(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 February 2001 (22.02.2001)

PCT

(10) International Publication Number
WO 01/12655 A1

(51) International Patent Classification[7]: C07K 7/04, A61K 38/10, A61P 35/00

(21) International Application Number: PCT/SE00/01576

(22) International Filing Date: 14 August 2000 (14.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9902899-5 13 August 1999 (13.08.1999) SE

(71) Applicant *(for all designated States except US)*: **KAROLINSKA INNOVATIONS AB** [SE/SE]; S-171 77 Stockholm (SE).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **TOFTGÅRD, Rune** [SE/SE]; Sätragårdsvägen 209, S-127 36 Skärholmen (SE).

(74) Agents: **BERG, S., A.** et al.; Albihns Patentbyrå Stockholm AB, P.O. Box 5581, S-114 85 Stockholm (SE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/12655 A1

(54) Title: PEPTIDES CONSISTING OF FRAGMENTS OF GLI-1 AND SUFUH AND THEIR USE

(57) Abstract: The present invention relates to the field of mammalian signalling pathways, and more precisely to the physical interaction between two components of the Sonic hedgehog (Shh)-Patched (Ptch) signalling pathway, namely GLI-1 and SUFUH. The invention provides peptides consisting of fragments of GLI-1 and SUFUH, respectively which are able to specifically bind to SUFUH and GLI-1, respectively. The invention also provides monoclonal antibodies and antibody fragments specifically binding to these peptides, as well as pharmaceutical compositions containing the peptides, antibodies and/or antibody fragments, said pharmaceutical compositions being useful for treating cancer and diseases influencing cell differentiation and tissue development.

New peptides, corresponding antibodies, and compositions thereof.

The present patent application relates to the field of mammalian signalling pathways, and more precisely to the physical interaction between two components of the Sonic hedgehog (Shh)-Patched (Ptch) signalling pathway, namely GLI-1 and SUFUH.

The Sonic hedgehog (Shh) – Patched (Ptch) signalling pathway is of key importance for both normal development and carcinogenesis, as shown by the presence of mutations in genes encoding components of this pathway in human malformation and cancer-prediposing syndromes (holoprosencephaly [Roessler et al., Hum. Mol. Genet. **6**:1847-1853 (1997)], nervoid basal cell carcinoma syndrome [Hahn et al., Cell **85**:841-851 (1996); Johnson et al., Science **272**:1668-1671 (1996)] and Greig cephaloplysyndactyly [Vortkamp et al., Nature **352**:539-540 (1991)].

Intracellular transduction of the Hedgehog signal in *Drosophila* involves derepression of the seven-transmembrane-domain protein Smoothened and release of a large complex containing Costal-2 (Cos-2), Fused (Fu), Suppressor-of-Fused (Su(fu)) and the transcriptional effector Cubitus interruptus (Ci) from microtubules at which they are anchored, as well as blockage of the proteolytic conversion of full-length Ci to its amino-terminal repressor form [Ruiz i Altaba, Cell **90**:193-193 (1997); Monnier et al., Curr. Biol. **8**:583-586 (1998); Methot et al., Cell **96**:819-831 (1999)]. In Drosophila only the repressor form of Ci is detected in the nucleus indicating that a specific mechanism(s) may prevent nuclear entry of full-length Ci. A separate activation step has been implicated in the conversion of full-length Ci to a transcriptional activator [Methot et al., Cell **96**:819-831 (1999); Ohlmeyer et al., Nature **396**:749-753 (1998)].

In mammals, functions corresponding to that of Ci have been assigned to the zinc-finger-containing and DNA-binding proteins Gli 1-3, which are expressed in an overlapping pattern adjacent to cells secreting Shh or the homologous Indian hedgehog

(Ihh) and Desert hedgehog (Dhh). *GLI-3* mutations detected in human disorders result in expression of truncated GLI-3 proteins that mimic natural Ci processing with respect to their altered subcellular localization and transactivation properties in HeLa cells (the full-length protein is cytoplasmic)[Shin et al., Proc.Natl.Acad.Sci. USA **96**:2880-2884 (1999)]. In the case of GLI-1, initial studies using the D259MG glioma cell line, which contains an amplified *GLI-1* locus, showed that this protein is nuclear in localization [Kinzler et al., Mol.Cell.Biol. **10**:634-642 (1990)]; nuclear GLI-1 localization was also seen after *GLI-1* cDNA was transfected into COS cells [Dahmane et al., Nature **389**:876-881 (1997)]. In contrast, GLI-1 in human basal cell carcinomas (BCCs) was again cytoplasmic [Dahmane et al., Nature **389**:876-881 (1997)]. Most recently it has been shown that, upon overexpression, all three GLI proteins may show either cytoplasmic or nuclear localization, depending on the cellular context [Ruiz i Altaba, Development **126**:3205-3216 (1999)]. Analysis of the proteolytic processing of vertebrae Gli proteins using mouse embryo extracts has shown the appearance of shorter variants of endogenous Gli-3, but not of Gli-1 [Dai et al., J. Biol. Chem. **274**:8143-8152 (1999)], whereas after overexpression in frog embryos shorter variants of all three Gli proteins were observed [Ruiz i Altaba, Development **126**:3205-3216 (1999)]. However, as overexpressed full-length Gli proteins can be detected in the nucleus, proteolytic processing does not appear to be necessary for nuclear import. Taken together, these data strongly indicate that, in mammalian cells, there may be mechanisms that regulate the subcellular localization of both full-length and processed Gli proteins. Such mechanisms could involve the regulation of interactions of the Gli proteins with anchoring proteins, or the modification of the Gli proteins themselves, including proteolytic processing.

The above mentioned component of this system, the Suppressor of Fused, Su(fu), in *Drosophila* [Pham et al., Genetics **140**:587-598 (1995)], has sofar not been been identified in man. In insects, Su(Fu) has been shown to form a complex together with Costal-2 (Cos2) and the transcription factor Cubitus interruptus (Ci) [Monnier et al., Cur-

rent Biol. 8:583-586 (1998)] which all are involved in the transduction of the Hedgehog signal.

The hedgehog signalling pathway is influenced by numerous autoregulatory and other control mechanisms as for example transcriptional regulation of Patched and Gli by Gli itself [Alexandre et al., Genes Dev. **10**:2003-2013 (1996)], and proteolytic generation of Ci and Gli isoforms with opposite transactivation properties [Jiang et al., Nature **391**:493-496 (1998); Ohlmeyer et al, Nature **396**:749-753 (1998); Methot et al., Cell **96**:819-831 (1999)]. Therefore, it is reasonable to assume that Su(fu) activity would also be controlled at various levels.

Accordingly, it would be highly desirable to find the control mechanisms behind the Sonic Hedgehog (Shh) – Patched (Ptch) signalling pathway, and to be able to use them as a basis for a pharmaceutical preparation for treating cancer and diseases influencing cell differentiation and tissue development.

Summary of the Invention

The present invention relates to the field of mammalian signalling pathways, and more precisely to the physical interaction between two components of the Sonic hedgehog (Shh)-Patched (Ptch) signalling pathway, namely GLI-1 and SUFUH. The invention provides peptides consisting of fragments of GLI-1 and SUFUH, respectivelwhich are able to specifically bind to SUFUH and GLI-1, respectively. The invention also provides monoclonal antibodies and antibody fragments specifically binding to these peptides, as well as pharmaceutical compositions containing the peptides, antibodies and/or antibody fragments, said pharmaceutical compositions being useful for treating cancer and diseases influencing cell differentiation and tissue development.

Detailed description of the invention

As already mentioned, the present invention relates to the physical interaction between GLI-1 and SUFUH. Accordingly, the present invention provides polypeptides derived from GLI-1 and SUFUH, respectively, which polypeptides show affinity to SUFUH and GLI-1, respectively.

Nucleic acid sequences, peptides and proteins are referred to in the description, claims and sequence listing by the following abbreviations:

SEQ.ID.NO.1 discloses the amino acid sequence of SUFUH. This sequence has also been disclosed in WO 99/32517.

SEQ.ID.NO.2 and SEQ.ID.NO.3 disclose sequences corresponding to amino acid residues 380 – 484 and 387 – 484, respectively, of SUFUH. These fragments comprise a GLI-1 binding site.

SEQ.ID.NO.4 discloses a sequence corresponding to amino acids 387 – 405 of SUFUH. Peptides comprising this sequence have been shown to physically interact with GLI-1.

SEQ.ID.NO.5 presents the sequence of GLI-1. It is disclosed in Kinzler et al., Nature **332**:371-4.

SEQ.ID.NO.6 discloses a sequence corresponding to amino acid residues 1 – 407 of GLI-1. This fragment comprises sites physically interacting with SUFUH.

SEQ.ID.NO.7 discloses a sequence corresponding to amino acids 381 – 407 of GLI-1. This fragment comprises sites controlling shuttling of GLI-1 and is blocked when SUFUH interacts with this protein.

SEQ.ID.NO.8 presents a sequence corresponding to amino acids 62 – 82 of GLI-1. This fragment comprises sites binding to SUFUH.

SEQ.ID.NO.9 shows the amino acid sequence of SUFU from *Drosophila melanogaster*.

SEQ.ID.NO.10 presents the amino acid sequence of SUFUHΔ, a SUFUH variant that has been truncated in the carboxy terminal after amino acid 385.

SEQ.ID.NO.11 shows the amino acid sequence of haematoglutinin-tagged GLI-1.

SEQ.ID.NO.12 shows the amino acid sequence of myc-tagged SUFUH.

SEQ.ID.NO.13 shows the amino acid sequence of a haematoglutinin-tagged fragment comprising amino acids 1-407 of GLI-1

SEQ.ID.NO.14 shows the amino acid sequence of myc-tagged SUFUHΔ.

SEQ.ID.NO.15 shows the amino acid sequence of an amino-terminal fragment of GLI-1 comprising amino acids 1 – 525.

SEQ.ID.NO.16 shows the amino acid sequence of an amino-terminal fragment of GLI-1 comprising amino acids 1 – 525 in which leucines 495 and 501 have been altered to alanines.

SEQ.ID.NO.17 shows the amino acid sequence of SU(FU)-XL.

SEQ.ID.NO.18 shows the amino acid sequence of SU(FU)-LK.

SEQ.ID.NO.19 shows the amino acid sequence of myc-tagged SU(FU)-Tt.

SEQ.ID.NO.20 shows the amino acid sequence of myc-tagged SU(FU)-XL.

SEQ.ID.NO.21 shows the amino acid sequence of myc-tagged SU(FU)-LK.

SEQ.ID.NO.22 shows the amino acid sequence of myc-tagged SU(FU)-Tt.

SEQ.ID.NO.23 shows the DNA sequence of SUFUH.

SEQ.ID.NO.24 shows the DNA sequence of SUFUH-XL.

SEQ.ID.NO.25 shows the DNA sequence of SUFUH-LK.

SEQ.ID.NO.26 shows the DNA sequence of SUFUH-Tt.

SEQ.ID.NO.27 shows the DNA sequence of GLI-1.

SEQ.ID.NOs.28-31 show the DNA sequence of four oligonucleotide primers useful in the chromosomal localization of SUFUH.

The polypeptides of the invention have one or more of the following properties. In a first embodiment, the polypeptides comprise fragments and/or derivatives of SUFUH that interacts with GLI-1. In a second embodiment, the polypeptides comprises fragments and/or derivatives of GLI-1 that binds to SUFUH. By contacting GLI-1 and SUFUH with the polypeptides of the invention *in vivo*, it should be possible to affect the Sonic hedgehog (Shh) – Patched (Ptch) signalling pathway which could be useful in treatment of cancer.

The nucleic acid sequences of the present invention are preferably DNA, though they may be RNA. Nucleic acid sequences of the invention will typically be in isolated or

substantially isolated form. For example up to 80, up to 90, up to 95 or up to 100% of the nucleic acid material ir a preparation of a nucleic acid of the invention will typically be nucleic acid according to the invention. Alterations, isolations or syntheses of the nucleic acid sequences of the invention may be performed by any conventional method, for example by the methods of Sambrook et al (Molecular Cloning: A Laboratory Manual; 1989).

A substitution, deletion or insertion may suitably involve one or more amino acids, typically from one to five, one to ten or one to twenty amino acids, for example, a substitution, deletion or insertion of one, two, three, four, five, eight, ten, fifteen, or twenty amino acids. Typically, a polypeptide of the invention has at least 70% at least 80%, at least 90%, or at least 95% sequence identity to the sequence of any one of SEQ. ID. Nos. 2 to 4, or SEQ.ID.Nos 6 to 8.

In general, the physicochemical nature of the sequence of SEQ. ID. Nos. 2 to 4 or SEQ.ID.Nos 6 to 8 should be preserved in a polypeptide of the invention. Such sequences will generally be similar in charge, hydrophobicity and size to that of SEQ. ID. Nos. 2 to 4 or SEQ.ID.Nos 6 to 8. Examples of substitutions that do not greatly affect the physicochemical nature of amino acid sequences are those in which an amino acid from one of the following groups is substituted by a different amino acid from the same group:

H, R and K

I, L, V and M

A, G, S and T

D, E, Q and N.

Where polypeptides of the invention are synthesised chemically, D-amino acids (which do not occur in nature) may be incorporated into the amino acid sequence at sites where they do not affect the polypeptides biological properties. This reduces the polypeptides' susceptibility to proteolysis by the recipient's proteases.

The nucleic acid sequences encoding the polypeptides of the invention may be extended at one or both ends by any non-wild-type sequence.

Thus, the polypeptides of the invention may be extended at either or both of the C- and N- termini by an amino acid sequence of any length. For example, an extension may comprise up to 5, up to 10, up to 20, up to 50, or up to 100 or 200 or more amino acids.

A polypeptide of the invention may be subjected to one or more chemical modifications, such as glycosylation, sulphation, COOH-amidation or acylation. In particular, polypeptides that are acetylated at the N-terminus are preferred, as are polypeptides having C-terminal amide groups. Preferred polypeptides may have one or more of these modifications. For example, particularly preferred peptides may have a C-terminal amide group and N-terminal acetylation.

A polypeptide of the invention may form part of a larger polypeptide comprising multiple copies of the sequence of one or more of SEQ. ID. Nos. 2 to 4 or SEQ.ID.Nos. 6 to 8 or a sequences related to them in any of the ways defined herein.

Polypeptides of the invention typically comprise at least 15 amino acids, for example 15 to 20, 20 to 50, 50 to 100 or 100 to 200 or 200 to 300 or 300 to 400 amino acids.

Polypeptides according to the invention may be purified or substantially purified. Such a polypeptide in substantially purified form will generally form part of a preparation in which more than 90%, for example up to 95%, up to 98% or up to 99% of the peptide material in the preparation is that of a polypeptide or polypeptides according to the in-

vention.

A nucleic acid sequence according to the invention may be included within a vector, suitably a replicable vector, for instance a replicable expression vector.

A replicable expression vector comprises an origin of replication so that the vector can be replicated in a host cell such as a bacterial host cell. A suitable vector will also typically comprise the following elements, usually in a 5' to 3' arrangement: a promoter for directing expression of the nucleic acid sequence and optionally a regulator of the promoter, a translational start codon and a nucleic acid sequence according to the invention encoding a polypeptide having the ability to either bind to GLI-1 or SUFUH. A non-replicable vector lacks a suitable origin at replication whilst a non-expression vector lacks an effective promoter.

The vector may also comprise one or more non-coding sequences 3' to the sequence encoding the polypeptide of the invention. These may be from *H. sapiens* (the organism from which the sequences of the invention are derived) or the host organism which is to be transformed with the vector or from another organism.

In an expression vector, the nucleic acid sequence of the invention is operably linked to a promoter capable of expressing the sequence. "Operably linked" refers to a juxtaposition wherein the promoter and the nucleic acid sequence encoding the polypeptide of the invention are in a relationship permitting the coding sequence to be expressed under the control of the promoter.

The vector may be of any type. The vector may be in linear or circular form. For example, the vector may be a plasmid vector. Those of skill in the art will be able to prepare suitable vectors comprising nucleic acid sequences encoding polypeptides of the invention starting with widely available vectors which will be modified by genetic

engineering techniques such as those described by Sambrook et al (Molecular Cloning: A Laboratory Manual; 1989).

In an expression vector, any promoter capable of directing expression of a sequence of the invention in a host cell may be operably linked to the nucleic acid sequence of the invention.

Such vectors may be used to transfect or transform a host cell. Depending on the type of vector, they may be used as cloning vectors to amplify DNA sequences according to the invention or to express this DNA in a host cell.

A further embodiment of the invention provides host cells harbouring vectors of the invention, i.e. cells transformed or transfected with vectors for the replication and/or expression of nucleic acid sequences according to the invention. The cells will be chosen to be compatible with the vector and may for example be bacterial cells. Transformed or transfected bacterial cells, for example *E. coli* cells, will be particularly useful for amplifying nucleic acid sequences of the invention as well as for expressing them as polypeptides.

The cells may be transformed or transfected by any suitable method, such as the methods described by Sambrook et al (Molecular cloning: A Laboratory Manual; 1989). For example, vectors comprising nucleic acid sequences according to the invention may be packaged into infectious viral particles, such as retroviral particles. The constructs may also be introduced, for example, by electroporation, calcium phosphate precipitation, biolistic methods or by contacting naked nucleic acid vectors with the cells in solution.

In the said nucleic acid vectors with which the host cells are transformed or transfected, the nucleic may be DNA or RNA, preferably DNA.

The vectors with which the host cells are transformed or transfected may be of any suitable type. The vectors may be able to effect integration of nucleic acid sequences of the invention into the host cell genome or they may remain free in the cytoplasm. For example, the vector used for transformation may be an expression vector as defined herein.

The present invention also provides a process of producing polypeptides according to the invention. Such a process will typically comprise transforming or transfecting host cells with vectors comprising nucleic acid sequences according to the invention and expressing the nucleic acid sequence in these cells. In this case, the nucleic acid sequence will be operably linked to a promoter capable of directing its expression in the host cell. Desirably, such a promoter will be a "strong" promoter capable of achieving high levels of expression in the host cell. It may be desirable to overexpress the polypaptide according to the invention in the host cell. Suitable host cells for this purpose include yeast cells and bacterial cells, for example *E. coli*. However, other expression systems can also be used, for example baculovirus systems in which the vector is a baculovirus having in its genome nucleic acid encoding a polypeptide of the invention and expression occurs when the baculovirus is allowed to infect insect cells.

The thus produced polypeptide of the invention may he recovered by any suitable method known in the art. Optionally, the thus recovered polypeptide may be purified by any suitable method, for example a method according to Sambrook et al (Molecular Cloning: A Laboratory Manual).

The polypeptides of the invention may also be synthesised chemically using standard techniques of peptide synthesis. For shorter polypeptides, chemical synthesis may be preferable to recombinant expression. In particular, peptides of up to 20 or up to 40 amino acid residues in length may desirably be synthesised chemically.

The present invention also provides antibodies to the polypeptides of the invention.

These antibodies may be monoclonal or polyclonal. For the purposes of this invention, the term "antibody", includes fragments of whole antibodies which retain their binding activity for a target antigen. Such fragments include Fv, F(ab') and F(ab').sub.2 fragments, as well as single chain antibodies.

The antibodies may be produced by any method known in the art, such as the methods of Sambrook et al (Molecular Cloning: A Laboratory Manual; 1989). For example, they may be prepared by conventional hybridoma techniques or, in the case of modified antibodies or fragments, by recombinant DNA technology, for example by the expression in a suitable host vector of a DNA construct encoding the modified antibody or fragment operably linked to a promoter. Suitable host cells include bacterial (for example *E. coli*), yeast, insect and mammalian cells. Polyclonal antibodies may also be prepared by conventional means which comprise inoculating a host animal, for example a rat or a rabbit, with a peptide of the invention and recovering immune serum.

The present invention also provides pharmaceutical compositions comprising polypeptides or antibodies of the invention. In a first embodiment, the compositions comprise antibodies, fragments and/or derivatives of SUFUH that interacts with GLI-1. In a second embodiment, the compositions comprises antibodies, fragments and/or derivatives of GLI-1 that binds to SUFUH. By administrating pharmaceutical compositions of the present invention to a patient in need thereof, it should be possible to affect the Sonic hedgehog (Shh) – Patched (Ptch) signalling pathway which could be useful in treatment of cancer.

The compositions of the invention may be administered to mammals including humans by any route appropriate. Suitable routes include topical application in the mouth, oral delivery by means of tablets or capsule and parenteral delivery, including subcutaneous, intramuscular, intravenous and intradermal delivery. Preferred routes of administration are injection, typically subcutaneous or intramuscular injection.

The compositions of the invention may be administered to the subject alone or in a liposome or associated with other delivery molecules. The effective dosage depends on many factors, such as whether a delivery molecule is used, the route of delivery and the size of the mammal being treated. Typical doses are from 0.1 to 100 mg of the polypeptide of the invention per day, for example 0.1 to 1 mg, and 1 to 5 mg, 5 to 10 mg and 10 to 100 mg per dose. Dosage schedules will vary according to, for example, the route of administration, the species of the recipient and the condition of the recipient.

While it is possible for polypeptides of the invention to be administered alone it is preferable to present them as pharmaceutical formulations. The formulations of the present invention comprise at least one active ingredient, a polypeptide or an antibody of the invention, together with one or more acceptable carriers thereof and optionally other therapeutic ingredients. The carrier or carriers must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not deleterious to the recipients thereof, for example, liposomes.

Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostatis, bactericidal antibiotics and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents, and liposomes or other microparticulate systems which are designed to target the compound to blood components or one or more organs.

Of the possible formulations, sterile pyrogen-free aqueous and non-aqueous solutions are preferred. Also preferred are formulations in which the polypeptides of the invention are contained in liposomes. Injection solutions and suspensions may be prepared extemporaneously from sterile powders, granules and tablets of the kind previously described.

It should be understood that in addition to the ingredients particularly mentioned above the formulations of this invention may include other agents conventional in the art having regard to the type of formulation in question.

We have described a human homologue of *Drosophila* Su(fu) and presented strong evidence that the function of this protein is highly conserved. Our results are consistent with, and provide a mechanistic explanation for, the genetic data obtained from studies of *Drosophila* that indicate an inhibitory role for Su(fu) in the Hedgehog signalling pathway [Ingham et al., EMBO J. **17**:3505-3511 (1998)]. We have shown, for the first time, to our knowledge, that SUFUH inhibits GLI-mediated transcription and the cellular response to SHH, and our results implicate the nucleus as a new intracellular site of action for SUFUH.

Moreover, we have shown that active nuclear export is a central mechanism for exclusion of full-length GLI-1 from the nucleus, offering a plausible explanation for the observation in *Drosophila* that Ci remains cytoplasmic in cells devoid of either Su(fu) or Cos2 [Monnier et al., Curr. Biol. **8**:583-586 (1998); Ingham et al., EMBO J. **17**:3505-3511 (1998)]. Interestingly, it has been proposed [Methot et al., Cell **96**: 819-831 (1999); Ohlmeyer et al., Nature **396**:749-753 (1998)] that a separate, uncharacterized activation step is required to convert full-length Ci to a transcriptional activator and it is tempting to speculate that this step involves regulation of nuclear-cytoplasmic shuttling. Further support for the biological relevance of such a regulated event comes from studies showing that GLI proteins show different subcellular distributions in different cell types [Shin et al., Proc. Natl. Acad. Sci. USA **96**:2880-2884 (1999); Ruiz i Altaba et al., Development **126**:3205-3216 (1999)], and that for all analysed artificially truncated or proteolytically processed Ci or GLI 1-3 proteins, subcellular distribution is consistent with inclusion or exclusion of predicted NES motifs [Shin et al., Proc. Natl. Acad. Sci. USA **96**:2880-2884 (1999); Ruiz i Altaba et al., Development **126**:3205-3216 (1999); Dai et al., J. Biol. Chem. **274**:8143-8152 (1999); Aza-Blanc et al., Cell **89**:1043-1053 (1997)] (Fig. 6b). It should be noted that in the two cases in which loca-

lization has been determined *in vivo* without vas overexpression (that is, for full-length Ci in the fly [Aza-Blanc et al., Cell **89**:1043-1053 (1997)] and for GLI-1 in human BCCs [Dahmane et al., Nature **389**:876-881], the localization has been primarily cytoplasmic. A likely scenario, taking into account the role of GLI-1 as a transforming oncoprotein and a cell-fate determinant that responds to graded levels of SHH signalling, involves a finely tuned balance of activated nuclear entry and nuclear export.

The retention of GLI-1 in the cytoplasm by SUFUH when nuclear export is compromised, and the similar SUFUH-mediated retention in the cytoplasm of an otherwise constitutively nuclear truncated GLI-1 variant that lacks the NES, indicates that SUFUH could block nuclear enty of GLI-1, possibly by masking a nuclearlocalization signal, and thereby inhibit transcriptional activation of target genes (Fig. 8). Consistent with this idea, a truncated SUFUH variant unable to repress GLI1-induced transcriptional activation is also unable to modify the subcellular localization of GLI-1. Whether or not binding of SUFUH to GLI-1 on DNA, or elsewhere in the nuclear compartment, actually acts to repress of block activation of transcription, alone or in combination with cytoplasmic retention of GLI-1, remains an interesting question for future studies. The expression of Sufu in cells next to Shh- or Ihh-producing cells during mouse and human embryogenesis, coupled with the ability of Sufu to inhibit Gli-mediated transcriptional activation, indicates that an important function of Sufu may be to act in an intracellular negative feedback mechanism and to impose thresholds on the responsiveness of cells to Shh and Ihh. A similar role for D-Axin has been proposed as regards Wingless signalling in *Drosophila* [Hamada et al., Science **283**:1739-1742 (1999)].

Finally, GLI-1 is an oncoprotein that is able to transform cells *in vitro* in corporation with E1A [Ruppert et al., Mol. Cell Biol. **11**:1724-1728 (1991)]; overexpression of GLI-1 induces epidermal tumours in frogs and consistent overexpression of GLI-1 is observed in BCCs [Dahmane et al., Nature **389**:876-881 (1997)]. This possibility is now being investigated in both human and murine systems.

We have also found two organ-specific expression variants of SU(FU) which might provide a further regulatory mechanism. Both these variants lack the ability of binding GLI1 thereby uncoupling the transcription factor from one of its cytoplasmic retention mechanisms. The sources of both variants, testis and leukocytes, are proliferating tissues. Since in both cases the variants contributed only partially to the total amount of SU(FU) expressed, it might be a particular subpopulation of cells which expresses them. Since we have found that SU(FU) is expressed in all adult tissues with comparable abundance, it is possible that in the adult the activity of SU(FU) is not primarily regulated by the level of SU(FU) transcription but rather by the exon composition. Another transcriptional variation has been described [Stone et al., J. Cell Science **112**:4437-4448 (1999)] which is expressed dominantly in testis and has the N-terminal 432 amino acids in common with the standard SU(FU) but then a STOP codon after one further altered amino acid. According to our data this variant comprises exons 1-10.

In the embryo, Su(fu) was expressed with tissue-specific preference. As stated above, the sites of expression grossly congrue with the sites of expression of other members of the hedgehog signalling pathway, in particular with the expression of Gli3. However, in detail this overlap is not found in all cases, and it is not always complete. In some cases the expression of Su(fu) is more restricted than the expression of Gli1. For example in mouse skin, until dpc 15.5, Su(fu) is expressed only in the vibrissae follicles stronger than the general background while Gli1 is expressed in vibrissae and pelagic hair follicles from their onset [St-Jaques et al., Curr. Biol. **8**:1058-1068 (1998)]. In other cases the expression of Gli1 is more restricted than the expression of Su(fu). For example, Gli1 [Hynes et al., Neuron **19**:15-26 (1997)] and Gli2 [Matise et al., Development **125**:2759-2770 (1998)] were reported to be expressed only in parts of the brain and the neural tube. In contrast, Su(fu) was expressed in the entire brain and spinal cord. Furthermore, we found examples where predominant Su(fu) and Gli expression sites are adjacent to each other (Gli1 was shown to be expressed in the papillal

epidermis of the tongue [Hall et al., J. Comp. Neurol. **406**:143-155 (1999)], while Su(fu) was expressed in the tongue mesenchyme). Since Gli1 activity is negatively controlled by Su(fu) and Gli1 expression is a positive target of Gli1 activity the expression of Gli1 could be modulated by the expression of Su(fu). By inducing the transcription both of Ptch1 and of itself, Gli1 triggers a negative and a positive feedback cycle. We do not know which of these activities would dominate under physiological conditions. Therefore we cannot anticipate that Su(fu) should always have an inhibitory effect on hedgehog signalling. Furthermore, it is possible that Gli1 could influence the expression pattern of the other types of Gli at neighbouring sites. One such example was observed in the digits where Gli1 and Su(fu) are expressed in the condensing mesenchyme while Gli2 and Gli3 are expressed in the interdigital mesenchyme [Mo et al., Development **124**:113-123 (1997)]. However, only in Gli2 and Gli3 mutant mice the development of the long and short bones of the limbs is affected [Mo et al., Development **124**:113-123 (1997)] while in Gli1 mutant mice (deletion of the zinc finger region) development appears to be normal [Park et al., Development **127**:1593-1605 (2000)].

We have investigated the possibility that the mammalian SU(FU) and Su(fu) would be the genes responsible for the SHFM3 and Dac disorders. During the course of these studies the Dac mutation was identified as an alteration in the previously unknown F-Box/WD40 gene *dactylin* [Sidow et al., Nat. Genet. **23**:104-107 (1999)]. At the same time, the human homolog for *dactylin* was also published [Ianakev et al., Biochem. Biophys. Res. Commun. **261**:64-70 (1999)]. For the human *Dactylin*, however, a link to SHFM3 is not yet demonstrated.

A role of *SU(FU)* in *SHFM3* still appears as an intriguing possibility not only because of a possible functional relation of dactylin and Su(fu) and because of the overlapping of their chromosomal loci but also because SU(FU) and other members of the hedgehog signalling pathway are expressed at the ossification zone of the digits in mice (this report) and humans. Moreover, other members of the hedgehog signalling pathway

were shown to be involved in anomalies of finger development. Preaxial polydactyly in mouse has been shown to be caused by ectopic expression of Indian hedgehog in the limbs which results in ectopic Gli1 expression [Yang et al., Development **125**:3123-3132 (1998)]. A loss of function-mutation in *GLI3*, associated with an increased Sonic hedgehog expression, has been shown to cause Postaxial Polydactyly Type A1 (41)[Radhakrishna et al., Am. J. Hum. Genet. **56**:597-604 (1997)] and Extratoes (Xt), a similar disorder, in the mouse [Hui et al., Nat. Genet. **3**:241-246 (1993)]. Other hereditary syndroms with a polydactyly phenotype in which *GLI3* is mutated are Greigs cephalopolysyndactyly [Vortkamp et al., Nature **352**:539-540 (1991)] and Pallister-Hall syndrome [Jang et al., J. Med. Genet. **34**:441-446 (1997)]. Thus, increased hedgehog signalling appears to correlate well with poly- and syndactyly. A decrease in digit number may conversely be associated with a repression of hedgehog signalling. Su(fu) is an antagonist of both Gli1 and Gli3 and gain of function mutations in Su(fu) would further reduce intracellular transduction of a hedgehog signal. As discussed by Özen et al., for all other known genes which are located in the same region (*FGF8*, *PAX2*, *ZNF32*, *FGFR2*, *HOX11*, *HMX2*, *WNT8b*, *LDB1*, *LBX1*, and *PITX1*) identity with the SHFM3 locus is unlikely for various reasons [Özen et al., Am. J. Hum. Genet. **64**:1646-1654 (1999)].

We have discovered a novel phenotype in the Dac mice which is the delay of ossification of that part of the skull that emerges under the membrane of the developing brain in direct vicinity to the neural tissue and which is similar to the human Adams-Oliver syndrome [Chitayat et al., Am. J. Med. Genet. **44**:562-566 (1992)]. Ossification here begins in wild-type mice at dpc 16.5 and has at dpc 17.5 already further progressed than in the Dac homozygous mouse at dpn 1 [Kaufman, The atlas of mouse development, Acad. Press Ltd. London, GB (1992)]. At dpc 15.5, the latest stage that we have investigated, Su(fu) was expressed strongly in the brain as it was during the entire period, but very weakly in the skin surrounding the brain.

It has been reported that Shh and Ptch1 are expressed in the osteogenic fronts of the calvarial bones [Kim et al., Development **125**:1241-1251 (1998)]. Expression was confined to the osteogenic fronts and later to the sutures between the parietal and between the parietal and the occipital bones, which are the bones affected in the Dac mouse, but not in the coronar suture which lays between affected and unaffected bones. This congrues very well with the assumption that a component of the hedgehog signalling pathway might be involved in causing the Dac phenotype. However, since no such calvarial phenotype has been reported for SHFM3 patients, SHFM3 and Dac may be similar but not identical, or the lack of this phenotype in humans may reflect species differences.

The invention will be further described with reference to the enclosed figures and tables, in which:

Table 1 shows exon-intron boundaries of the human *SU(FU)* and mouse *Su(fu)* genes;

Figure 1 presents the predicated amino-acid sequence (SEQ.ID.NO.1) encoded by the SUFUH gene and its expression in adult human tissues and mouse embryos.
a, The SUFUH amino-acid sequence (lower lines) is 40% identical to the *Drosophila melanogaster* Su(fu) sequence (SEQ.ID.NO.9)(upper lines). In the human sequence, the N-terminal part of the protein consists of a series of 20 short (A, G, P) and short hydroxylated (S, T) amino acids only (amino acids 6-25), including a duplicated PGPTAPPA motif (underlined), which is not found in the fly sequence. The region between amino acids 31 and 151 of SUFUH is highly conserved from fly to humans, with the C-terminal region exhibiting somewhat less similarity. A potential PEST motif in the *Drosophila* sequence (amino acids 308-327) [Pham et al., Genetics **140**:587-598 (1995)] is not conserved in the human sequence. Vertical lines between the two sequences indicate identical amino acids; two dots indicate small and one dot indicates large evolutionary distance between the amino acids.
b, A human multiple-tissue northern blot was probed with a fragment corresponding to base pairs, 1,182-1,630 of the human *SUFUH* cDNA. The sizes of the detected

mRNAs are ~ 5.5-6 and 2.8 kilobases (kb), respectively. An extra hybridizing band of 4.4 kb was observed in the heart only. 1, Heart; 2, brain; 3, placenta; 4, lung; 5, liver; 6, skeletal muscle; 7, kidney; 8, pancreas.

c, *Sufu* expression in the mouse embryo. Whole-mount *in situ* hybridization of an NMRI mouse embryo with a murine *Sufu* cDNA probe at d.p.c. 12.5; lateral view. The interfollicular tissue of the vibrissae field, the vertebral column and the interdigital mesenchyme are stained;

Figure 2 shows that SUFUH and PTCH1 are both highly expressed in osteoblasts in the perichondrium during human embryonic bone formation. Expression of the *SUFUH* and *PTCH1* mRNAs co-localize in osteoblasts in the perichondrium in the forming finger bones of a 12-week-old human embryo, as visualized by bright-field (**a-d**) and dark-field (**e**) photomicrography. **a**, An overview of the embryonic fingers counterstained with haematoxylin and eosin, **b**, A section of a finger hybridized with an antisense *SUFUH* mRNA probe, showing a strong autoradiographic signal (black grains) in the osteoblasts in the perichondrium (arrow) next to the chondrocytes (c). **c**, A serial section hybridized with a sense *SUFUH* probe showed no signal, **d**, **e**, Paired bright- and dark-field photomicrographs showing abdundant *PTCH1* mRNA signal in the same cellular compartment as shown in **b**. Scale bar in **a** represents 230 μm and in **b-e** represents 24 μm;

Figure 3 discloses that SUFUH inhibits transcriptional and biological activity of GLI-1. **a-c**, 293 or **d-f**, C3H10T1/2 cells were transiently transfected with a reporter construct and/or with indicated expression constructs. **a**, The reporter was a 4,3-kb fragment of the *PTCH1* promoter connected to the luciferase reporter gene. Expression of SUFUH alone has no effect on reporter gene activity whereas transfection of GLI-1 results in a five- to sixfold increase in luciferase activity. This activation can be inhibited by co-transfection of SUFUH in a dose-dependent manner. Values shown in parentheses are concentrations of *SUFUH*-expressing plasmid. **b**, Similar to the experiment shown in **a**, except that the reporter construct was a synthetic enhancer consisting

of 12 copies of the GLI-binding consensus site [Undén et al., Cancer Res. 57:2336-2340 (1997)]linked to the thymidine kinase (TK) basal promoter and the luciferase reporter gene. The results are qualitatively the same as those obtained with the *PTCH1* promoter except that GLI-1 activates this construct much more (200-300-fold). A corrected haemagglutinin-tagged GLI-1 construct (SEQ.ID.NO.11)(WTGLI1) activates the reporter even more strongly and this activation can be suppressed by coexpression of SUFUH but not by a construct containing the 385 N-terminal amino acids of SUFUH (SUFUHΔ)(SEQ.ID.NO10). **c**, SUFUH does not nonspecifically suppress transactivation by an unrelated transcription factor. A reporter containing two copies of a glucocorticoid response element connected to the TK basic promoter and luciferase reporter gene was used. The reporter is activated by a constitutively active glucocorticoid receptor (τlc) in which the DNA-binding domain is directly fused to the transactivation domain. **d**, Osteogenic differentiation of C3H10T1/2 cells transfected with GLI-1, detected by alkaline phosphatase staining. **e**, Example of a field of alkaline-phosphatase-negative cells (in this case transfected with both GLI-1 and SUFUH). **f**, Quantification of the differentiation assays, based on three independent experiments. The number of alkaline-phosphatase-positive cells following the transfection of different expression constructs is indicated. GLI-1 strongly induces osteogenic differentiation; this can be inhibited by full-length but not by truncated SUFUH. The N-terminal fragment of SHH also induces differentiation of C3H10T1/2 cells; this can be blocked by coexpression of SUFUH. The y-axis denotes the number of positive cells per transfected 35-mm well;

Figure 4 shows that SUFUH and GLI1 interact physically. The expression constructs used in this experiment were epitope-tagged (SUFUH with Myc and GLI-1 with haemagglutinin (HA)) and transfections to 293 cells were done as indicated. **a**, Full-length GLI-1 and SUFUH interact. The anti-Myc antibody does not precipitate HA-tagged GLI-1 when only HA-GLI1 (SEQ.ID.NO.11)(lane 1) or Myc-SUFUH (SEQ.ID.NO.12)(lane 2) is expressed in the cells. However, HA-GLI1 is already precipitated by the anti-Myc antibody in cells co-transfected with bort constructs (lane 3).

Conversely, Myc-SUFUH is not present in anti-HA precipitates from cells transfected with HA-GLI1 (lane 4) or Myc-SUFUH (lane 5) alone but is present when cells are co-transfected with both constructs (lane 6). IP, immunoprecipitation. Relative molecular mass standards (in thousands) are indicated to the sides of the gels. **b**, A C-terminally truncated GLI-1 construct containing the 407 N-terminal amino acids (GLI1(N407)(SEQ.ID.NO.13)) interacts with SUFUH, HA-GLI1(N407)(SEQ.ID.NO.13) co-precipitates with Myc-SUFUH (SEQ.ID.NO.12) only when they are co-expressed (lane 2). **c**, SUFUHΔ (SEQ.ID.NO.10) does not interact with GLI-1. Lane 1, an anti-Myc immunoprecipitation from cells expressing Myc-SUFUHΔ (SEQ.ID.NO.14) and HA-GLI1 (SEQ.ID.NO.11); lane 2, lysate containing both proteins. **d**, Both SUFUH (SEQ.ID.NO.1) and SUFUHΔ (SEQ.ID.NO.10) are produced at comparable levels and with the expected relative molecular masses. Lysates from cells transfected with Myc-SUFUH (SEQ.ID.NO.12) or Myc-SUFUHΔ (SEQ.ID.NO.14) were blotted with anti-Myc antibodies, **e**, SUFUH does not associate nonspecifically with expression constructs for IκBα or anti-Myc antibodies as indicated;

Figure 5 reveals that SUFUH but not SUFUHΔ retains nuclear GLI-1 variants in the cytoplasm. Subcellular localization of SUFUH and GLI-1 constructs was determined after transfection of epitope-tagged constructs into 293 cells. In each case, 450 ng of specific plasmid DNA was used, resulting in a 1:1 ratio in cotransfections. This ratio was shown previously (Fig. 3a, b) to cause strong transcriptional inhibition. **a**, SUFUH (SEQ.ID.NO.1) localizes to the cytoplasm and the nucleus, as determined by immunofluorescent microscopy. **b**, C-terminally truncated SUFUHΔ (SEQ.ID.NO.10) is also present in both cytoplasm and nucleus. **c**, EGFP-tagged GLI-1 shows cytoplasmic localization. **d**, EGFP-tagged GLI1(N407) is constitutively nuclear. **e**, **f**, EGFP-GLI1(N407) is partially retained in the cytoplasm by coexpression of SUFUH(**e**)(SEQ.ID.NO.1) but not SUFUHΔ(**f**) (SEQ.ID.NO.10). SUFUH was detected by indirect immunofluorescence; EGFP-GLI1 constructs were visualized di-

rectly (green); nuclei were stained with Hoechst dye (blue). The overlap of the two colours appears turquoise;

Figure 6 shows that GLI-1 accumulates in the nucleus when nuclear export is inhibited. **a**, A deletion construct containing the 525 N-terminal amino acids of GLI-1 (SEQ.ID.NO.15) is located in both the nucleus and the cytoplasm. **b**, The sequence of the putative NES in GLI-1 is conserved in all vertebrate species; the position of the first leucine in the protein sequence is indicated. Ci contains a putative NES in the same region (starting at amino acid 774) as well as two extra ones, at least one of which is conserved in vertebrate GLI-2 and GLI-3. **c**, A mutated GLI1(N525) (SEQ.ID.NO.16) protein, with leucines 495 and 501 mutated to alanines, localizes exclusively to the nucleus. **d**, Full-length GLI-1 (SEQ.ID.NO.5) translocates to the nucleus when nuclear export is inhibited by the addition of leptomycin B (LMB, 5 nM). **e**, Coexpression of SUFUH inhibits nuclear accumulation of GLI-1 in the presence of LMB. **f**, The C-terminally truncated SUFUHΔ (SEQ.ID.NO.10) construct has no effect on GLI-1 nuclear accumulation. The EGFP-GLI1 proteins appear green; the nuclei are stained blue with Hoechst dye; and the overlap between the two colours appears turquoise. For transfections, 405 ng specific plasmid DNA was used, resulting in a 1:1 ratio in co-transfections;

Figure 7 discloses that SUFUH can form a complex with constitutively nuclear GLI-1 variants bound to DNA. **a**, Nuclear GLI1(N407) (SEQ.ID.NO.6) expressed in 293 cells forms a specific complex with the consensus (labelled) GLI-binding oligonucleotide (lane 1); addition of excess unlabelled GLI oligonucleotide outcompetes the labelled oligonucleotide (lane 2). Coexpression of SUFUH results in complexes that are more retarded (lane 3); these complexes, by supershifting with anti-tag antibodies (anti-Myc for SUFUH and anti-EGFP for GLI1(N407), were shown to contain both GLI1(N407) (lane 4) and SUFUH (lane 5). Analysis of nuclear extracts after transfection of SUFUH alone revealed no retarded complex (late 6). **b**, Coexpression of SUFUH causes a redistribution of GLI1(N407) from the nuclear pellet to the nuclear exract and appear-

ance of GLI1(N407) in the cytoplasmic fraction. The cytoplasmic and nuclear extracts were prepared in parallel using cells from three plates for each condition as described [Ausubel et al., Current Protocols in Molecular Biology (John Wiley, New York 1987)], using 1,2 M KCl for nuclear extraction. The remaining nuclear pellet was solubilized in SDS sample buffer and GLI1(N407) was detected by western blotting using anti-EGFP antibodies (Clontech). 15 µg protein from cytoplasmic and nuclear extracts or 5 µl solubilized nuclear pellet was loaded in each well. This represents 3%, 15% and 2,5% of the total cytoplasmic extract, nuclear extract and nuclear pellet, respectively. Relative molecular mass standards (in thousands) are indicated at the right;

Figure 8 discloses a model for the role of SUFUH in the nuclear-cytoplasmic shuttling of GLI1. GLI-1 is normally kept in the cytoplasm by active nuclear export mediated by its NES. We propose that SUFUH acts to inhibit GLI1-mediated transcriptional activation and thereby modulate cellular responsiveness to SHH/IHH. This inhibition can occur by preventing nuclear access of GLI-1, either by anchoring GLI-1 in the cytoplasm or by masking its putative NLS. SUFUH may also interfere with GLI-1 function in the nucleus by binding to GLI-1 on the DNA;

Figure 9 relates to organization and structure of the mouse *Suppressor of fused* (*Su(fu)*) gene locus and cDNAs. A) The mouse *Su(fu)* gene as assembled from mapping and partial sequencing of BAC clones 17985 and 17986. Boxes indicate position of exons. Filled and unfilled boxes indicate coding and non-coding regions, respectively. Arrows indicate direction of the transcriptional unit for *Su(fu)* and the linked upstream gene for *actin-related protein 1 a-isoform* (*Arp1*) and the downstream gene *for ADP-ribosyl-like protein 3* (*Arl3*). Only the exon containing the ATG translational start of *Arp1* and the exon after the ATG-containing exon of *Arl3* have been mapped and sequenced. Double arrowheads indicate start and end of BAC clones. Restriction enzyme sites of a selected set that have been sequenced and/or mapped are shown. Not all sites of the indicated enzymes have been mapped. E, EcoRI; B, BamHI; H, HindIII; K, KpnI; Xb, XbaI; Sm, SmaI. B) Structure of the fully sequenced mouse *Su(fu)* EST cDNAs

1195307 (mammary gland), 513730 (testis), 1224813 (heart), and 963990 (mammary gland), from the IMAGE consortium collection. Numbers 1 through 12 indicate the *Su(fu)* exons drawn to scale for a full-length mRNA transcript. ATG and TAG indicate the translational start and stop, respectively. The thick line represent the structural composition of each of the four EST clones positioned relative to the hypotetical full-length transcript;

Figure 10 shows transcriptional variations of human SU(FU). A. RT-PCR of a SU(FU)-fragment containing the exons 5 to 9 from various tissues. Tt and Lk indicate the respective variations. The other minor bands in lane 3 and lane 4 represent PCR products which are not related to SU(FU). B. Same as in A, after EcoNI digestion.The lower band represents the EcoNI susceptible SU(FU)-XL variant (SEQ.ID.NO.17). Since the mobility of the EcoNI-digested fragment equals the mobility of the SU(FU)-Lk band (SEQ.ID.NO.18), lane 8 was omitted here. Lane 1, thymus; 2, prostate; 3, spleen; 4, testis; 5, colon; 6, ovary; 7, small intestine; 8, peripheral leukocytes. C. Sequence of the intron 8a in SU(FU)-Tt (SEQ.ID.NO.19). The letters in italics signify the flanking exons 8 and 9;

Figure 11 discloses that Myc-SU(FU)-Lk (SEQ.ID.NO.21) and myc-SU(FU)-Tt (SEQ.ID.NO.22) cannot coprecipitate HA-GLI1 (SEQ.ID.NO.11). Each 10 µg of HA-GLI1 (SEQ.ID.NO.11) and myc-SU(FU)-Lk (SEQ.ID.NO.21)(lane 1), myc-SU(FU)-Tt (SEQ.ID.NO.22)(lane 2), myc-SU(FU)-XL (SEQ.ID.NO.20)(lane 3) and myc-SU(FU) (SEQ.ID.NO.12)(lane 4) were transfected into 293 cells. Lysates were precipitated with anti-myc (A) and anti-HA (B) antibodies and blotted with anti-myc antiserum. The numbers to the left indicate the positions of the molecular weight marker (kD) for each blot. The arrowhead next to blot B indicates the position of the coprecipitated myc-SU(FU)-XL (SEQ.ID.NO.20) and myc-SU(FU) (SEQ.ID.NO.12);

Figure 12 shows whole mount in situ hybridization of mouse embryos with a Su(fu) antisense probe. A. dpc 8.5. Somites and the neural tube hybridize with the probe. B.

dpc 9.5 lateral view. f, forelimb; h, heart; m, mandibular arch, s, somites; *, the mesenchyme surrounding the prospective mouth cavity. C. dpc 9.5, lateral view, prior to clearance with glycerol. nt, neural tube. D. dpc 10.5, dorsal view. b, brain vesicles. E. dpc 10.5, dorsal view of the tail after removal of the skin. drg, dorsal root ganglia. F. dpc 9.5, sagittal view. Besides the heart, the dorsal aorta (da) and the intersegmentary arteries (is) are stained. G. dpc 13.5, sagittal view. Most prominent is the staining of the spinal cord (sc) and the brain (b). Mandible (m), tongue (t), genital tubercle (g), and the mesenchyme close to the vertebrae (v) are also stained. H-J Staining of the mesenchyme adjcent to the ossification zone of the digits in the hindlimb at dpc 13.5 (H), 14.5 (I), 15.5 (J). The lines point to the Su(fu) expressing areas which moves distally with time. K. dpc 14.5 sagittal section of the snout. At this point, mainly the tongue (t) expresses Su(fu). Less staining was detected at the distalmost part of the snout (arrow) and of the mandible. L. dpc 15.5, sagittal view of the dorsal part of the thorax. The lines point to the Su(fu) expressing tracheoles; and

Figure 13 relates to Alizarin-Alcian Blue staining of the skull of a dpn 0 Dac/Dac (A) and wild type (B) mouse. While the occipital (O) and the frontal bones (F) are ossified in both cases (red staining), the supraoccipital (S), intraparietal (I), and parietal (P) bones are ossified in wildtype only, while in the Dac/Dac mouse there is poorly segmented cartilagenous mass (blue). Note also that the curvature of the skull is much flatter in the Dac/Dac mouse.

The invention will now be further disclosed in the following experimental part.

Experimental part

General Methods relating to Examples 1 - 8:

Cloning of mammalian SUFUH

ESTs for the human and mouse *Su(fu)* homologues were obtained from Genome Systems. The nucleotide sequence of the ESTs were determined and a full-length cDNA of the human *Su(fu)* homologue, referred to as *SUFUH*, was obtained by RACE, essentially as described [Zaphiropolous et al., DNA Cell. Biol. 15:1049-1056 (1996)].

Northern blot hybridization

Multiple tissue northern blots with human RNAs were obtained from Clontech. Fragments of the human *SUFUH* clone (SEQ.ID.NO.23) corresponding to base pairs 1,182-1,630 and to the3' untranslated region were labelled with ^{32}P using the Megaprime DNA-labelling kit (Amersham) and hybridized to these blots. After stripping the blots with 1% SDS, the amount and quality of the RNA was monitored with a probe specific for β-actin.

In situ hybridization

Digoxin-labelled riboprobes for murine *Sufu* were prepared using the DIG RNA-labelling kit (Boehringer Mannheim) using the mouse *Sufu* EST as a template and Sp6 and T7 poymerase to make sense and antisense probes, respectively. Timely mated NMRI mice were killed and the embryos collected, and, after fixation with 4% paraformaldehyde, whole-mount *in situ* hybridization was done according to standard procedures [Wilkinson, In Situ Hybridization: A Practical Approach (IRL, Oxford 1992). The hybridizing probe was detected using the DIG nucleic acid detection kit (Boehringer Mannheim). Unspecific binding of the probe was monitored by comparison of samples hybridized with sense and antisense probes.

Radioactive *in situ* hybridization was done as described [Undén et al., Cancer Res. 57:2336-2340 (1997)]. Briefly, a human *SUFUH* cDNA clone (SEQ.ID.NO.23) containing the complete coding region and a human *PTCH1* cDNA fragment (base pairs 190-628) were used to generate ^{35}S-labelled RNA probes. Sections were treated with

proteinase K and washed in 0,1 M triethanolamine buffer containing 0,25% acetic anhydride. Sections were hybridized with 2,5 x 10^6 c.p.m. of each of the labelled antisense and sense probes at 55°C. Autoradiographic exposure lasted for 2 days for *SUFUH* and 2 weeks for *PTCH1*. Formalin-fixed sections from a 12-week-old human embryo were obtained from the Department of Pediatric Pathology, Karolinska Hospital, Sweden.

Epitope-tagged expression constructs

Wild-type *GLI-1* was generated by PCR mutagenesis using oligonucleotides leading to a sequence change of G to A position 2,876 of the *GLI-1* cDNA (SEQ.ID.NO.27) [Hynes, Neuron **19**:15-26 (1997)]. Using the full-length cDNA clones encoding SUFUH, a variant of SUFUH truncated after amino acid 385 (SUFUHΔ)(SEQ.ID.NO.10) and human wild-type GLI-1 (SEQ.ID.NO.5) as templates, we generated plasmids coding for SUFUH (SEQ.ID.NO.1) and SUFUHΔ (SEQ.ID.NO.10) with an N-terminal Myc-epitope tag [Evan et al., Mol. Cell. Biol. **5**:3610-3616 (1985)], and plasmid encoding human GLI-1 with an N-terminal haemagglutinin tag (SEQ.ID.NO.11) [Pate, Gene **114**:285-288 (1992)], after ligation into the CMV-5 expression vector. *GLI-1* was subcloned into the EGFP-C3 vector (Clontech) for localization studies. GLI1(N407)(SEQ.ID.NO.6) was constructed by digesting the *GLI-1* expression vector with *Xcm*I and *Xba*I and GLI1(N525)(SEQ.ID.NO.15) by digesting the *GLI-1* expression vector with *Kpn*I and *Xba*I restriction enzymes followed by religation after blunt-ending. Leucines 495 and 501 were mutated to alanine in GLI(N525)(SEQ.ID.NO.16) by PCR-mediated mutagenesis (the first two nucleotides in codons 495 and 501 were mutated from CT to GC, resulting in the indicated amino-acid change). All PCR-general constructs were verified by sequencing.

Transient transfections

Human 293 cells (transformed kidney epithelial cell line) and C3H 10T1/2 cells (murine fibroblast cell line) were obtained from ATCC and grown according to the supplier's specifications. Cells were passaged to plates the day before the experiments and transfections were carried out with DNA complexed to the Superfect transfection reagent (Qiagen) according to the manufacturer's instruction. Expression was analysed 24 h, and differentiation 96 h, after transfections.

Reporter-gene and differentiation assays

The *PTCH-luc* construct [Shin et al., Proc. Natl. Acad. Sci. USA **96**:2880-2884 (1999)] and the *2Gluc-RE-TKO-luc* construct activable by a constitutively active glucocorticoid receptor, τlc [Almlöf et al., Mol. Cell. Biol. **17**:934-945 (1997)] have been described. The *12GLI-RE-TKO-luc* construct was generated by ligating annealed oligonucleotides corresponding to *GLI* consensus sequence into the *TKO-luc* vector. Luciferase assays were carried out as described construct [Shin et al., Proc. Natl. Acad. Sci. USA **96**:2880-2884 (1999)]. The results from at least four experiments from two independent transfections were compiled. The differentiation assays were done as described [Murone et al., curr. Biol. **9**:76-84 (1999)]. Briefly, 4 days after transfection of the indicated constructs, cells were fixed and stained for alkaline phosphatase activity with X-phos and pNBT (SIGMA) according to the manufacturer's instructions. The positive cells in each well were counted and at least three independent experiments were compiled for quantitative results.

Immunoprecipitations and western blotting

Lysates were prepared in immunoprecipitation buffer (PBS, 1% Nonidet-P40, protease inhibitors); immunoprecipitations were performed using 0,4 µg rabbit anti-haemagglutinin antibody and 20 µl packed protein G plus protein A agarose or 50 µl agarose-conjugated anti-Myc beads and separated along with the lysates on 7,5% SDS-PAGE, transferred to poly(vinylidene)fluoride (PVDF) membrane (Millipore), and in-

cubated with the indicated primary (Santa Cruz Biotechnology) and corresponding secondary horseradish-peroxidase-conjugated antibodies (Jackson Immunoresearch). Following washes, enhanced chemiluminescence detection (Pierce) was performed according to the manufacturer's instructions.

Localization studies

239 cells were transfected with appropriate expression constructs and, where indicated, 5 mM leptomycin B was added to the cells for the last 6 h. The cells were fixed for 10 – min in 2% formaldehyde, 0,2% glutaraldehyde and PBS, then permeabilized in PBSTX (PBS, 0,05% Triton-X100). Myc-SUFUH (SEQ.ID.NO.12) was detected by indirect immunofluorescent staining whereas EGFP constructs were detected directly by microscopy. Nuclei were stained with the Hoechst dye (1 $\mu g\ ml^{-1}$). At least 50 transfected cells were inspected in every case, all of which showed a similar staining pattern. For fluorescent microscopy we used a Leitz DMR confocal microscope (Leica); the images were captured with an ORBIS, cooled, chargecoupled-device camera (Spectrasource Instruments) and analysed using Slidebook software (Intelligent Imaging Innovations).

Subcellular fractionation and EMSAs

Nuclear and cytoplasmic extracts were prepared from cells transfected with either *SUFUH* or *GLI1(n407)* or both constructs, using standard protocols [Ausubel et al., Current Protocols in Molecular Biology (John Wiley, New York 1987)]. The nuclear extracts were incubated with end-labelled GLI-RE double-stranded oligonucleotides [Pham et al., Genetics **140**:587-598 (1995)] in the presence or absence of excess unlabelled oligonucleotides or anti-tag antibodies. The extracts were prepared and the complexes formed were resolved and detected as described [Ausubel et al., Current Protocols in Molecular Biology (John Wiley, New York 1987)].

Example 1: Cloning of human *SUFUH*.

By searching the expressed sequence tag (EST) database for genes homologous to the *Drosophila melanogaster Su(fu)*[Pham et al., Genetics **140**:587-598 (1995)], we found two clones encoding partial human and murine homologoues. Starting with the human EST, we cloned the complete coding region of the gene named *SUFUH* by rapid amplification of cDNA ends (RACE) technology. The resulting cDNA encompasses 2,239 base pairs and encodes a protein of 484 amino acids (SEQ.ID.NO.1). The SUFUH amino-acid sequence shows no striking homology to any other protein described so far, and shares 40% identity and 61% similarity with the *Drosophila* sequence (Fig. 1a)(SEQ.ID.NO.9). The region between amino acids 31 and 124 of SUFUH is highly conserved, being 66% identical to the *Drosophila* sequence, whereas the remaining parts have an identity of 32%. The partial murine Sufu amino-acid sequence is 98% identical to the human sequence.

Example 2: Expression of *SUFUH* in adult and embryonic tissues.

SUFUH (SEQ.ID.NO.1) was expressed in all adult human tissues tested (Fig. 1b). We detected two messenger RNAs, a minor, 2.8-kilobase, band and a major, 5.5-6-kilobase, band. Similar hybridization patterns were seen with probes corresponding to the coding region and to the 3'-untranslated region, indicating that the two mRNAs may be processing variants. The use of an alternative polyadenylation signals is also suggested by the identification of other mouse ESTs containing additional 3'-untranslated regions (our unpublished observations).

To test whether vertebrate Sufu is expressed in a pattern consistent with a potential role in mediating Shh signalling during embryogenesis, we used whole-mount in situ hybridization to analyse *Sufu* expression in mouse embryos at days 8.5 to 15.5 post-coitum (d.p.c.). Throughout the entire period we observed signals in the neural tube and, at the later stages, in its derivatives, brain and spinal cord (data not shown). The somites expressed *Sufu* at all stages; the vibrissae field stained positively for *Sufu* from

d.p.c. 12.5 and onwards, with the vibrissae themselves being spared. The *Sufu* expression pattern during limb-bud development appeared to be separated into two distinct phases, with strong homogeneous staining all over the limb buds being observed from their emergence at d.p.c. 9.5, whereas at d.p.c. 12.5 only the interdigital mesenchyme of the limbs stained positively (Fig. 1c). This expression pattern partially overlaps with the expression of Ptch and the Ci homologues Gli 1-3, and is compatible with a conserved role for Sufu in Shh signalling.

To substantiate this observation in more detail and in the human system, we analysed the expression of SUFUH and PTCH1 in the developing limb of a 12-week-old human embryo by radioactive *in situ* hybridization (Fig. 2a-e). The results show marked *SUFUH* expression in the osteoblasts of the perichondrium, where *PTCH1* is also highly expressed. These findings are consistent with earlier observations in the avian and murine systems, in which Ptch1 and Gli1 are highly expressed in the same type of cells in response to Ihh secretion by prehypertrophic chondrocytes [Vortkamp et al., Science 273:613-622 (1996); Vortkamp et al., Mech. Dev. 71:65-76 (1998)]. Taken together, these results show that *SUFUH* is preferentially expressed in cells that receive a Hedgehog signal, and indicate that, during embryogenesis, SUFUH may be co-regulated with PTCH1 and GLI1. However, further studies will be required to demonstrate this relationship unequivocally, as well as to show a direct regulation of SUFUH by Hedgehog signalling.

Example 3: SUFUH inhibits transcriptional activity of GLI-1.

In a more direct attemption to confirm a role for SUFUH in SHH-PTCH-GLI signalling, we used reporter gene assays to analyse the effect of co-transfection of SUFUH and GLI-1 on GLI1-activated transcription and biological responses. As a reporter we used the firefly luciferase gene, controlled either by the human *PTCH1* promoter containing two well-conserved GLI-binding consensus sites (Fig. 3a, *PTCH-luc*) or by a synthetic promoter consisting of 12 repeated GLI consensus sites fused to the thymidi-

ne kinase basic promoter (Fig. 3b, *12GLI-RETKO-luc*). Both reporters were activated by GLI-1, whereas SUFUH had no effect by itself. However, co-transfection of SUFUH efficiently suppressed activation of these reporter gene constructs by GLI-1. As the originally isolated *GLI-1* gene contained a potentially activating mutation [Hynes et al., Neuron **19**:15-26 (1997)], we introduced a reverse mutation into the *GLI1* cDNA (SEQ.ID.NO.27) [Kinzler et al., Nature **332**:371-374 (1988)]. This wild-type *GLI-1* construct also contains an N-terminal haemagglutinin epitope tag and lacks the 3'-untranslated region; it activates the *12GLI-RETKO-luc* reporter very strongly (Fig. 3b). This activation can again be suppressed by co-transfection of SUFUH but not by a carboxy-terminally truncated SUFUH lacking amino acids 386-484 (SUFUHΔ)(SEQ.ID.NO.10). The wild-type *GLI-1* construct, or a similar construct encoding the enhanced green fluorescent protein (EGFP) tag, was used in all subsequent experiments. Specificity of transcriptional inhibition by SUFUH was confirmed by the fact that SUFUH did not interfere with transcriptional activation of a reporter gene induced by co-transfection of a constitutively active glucocorticoid receptor (Fig. 3c).

Example 4: SUFUH inhibits differentiation in response to SHH signalling.

To determine whether SUFUH was also able to inhibit a relevant cellular response to SHH or to one of its downstream effectors, GLI-1, we used the C3H10T1/2-cell osteogenic-differentiation assay [Nakamura et al., Biochem. Biophys. Res. Commun. **237**:465-469 (1997); Murone et al., Curr. Biol. **9**:76-84 (1999)]. Transfection of GLI-1 induced differentiation (Fig. 3d-f), as measured by alkaline phosphatase staining. Co-transfection of SUFUH but not of SUFUHΔ (Fig. 3f) inhibited this response. SUFUH also inhibited the differentiation response in cells transfected (Fig. 3f) with the N-terminal active fragment of SHH.

Example 5: SUFUH and GLI-1 are present in an intracellular complex.

These results indicated that the functions of GLI-1 and SUFUH may also be closely linked in the mammalian signal-transduction pathway and raised the possibility that the two proteins might associate physically, as reported for the fly counterparts [Monnier et al., Curr. Biol. **8**:583-586 (1998)]. To test this possibility, we transiently expressed epitope-tagged GLI-1 and SUFUH, either alone or in combination, in 293 cells. Haemagglutinin-tagged GLI-1 (SEQ.ID.NO.11) co-precipitated with Myc-tagged SUFUH when the latter protein was immunoprecipitated with an anti-Myc antibody; conversely, when haemagglutinin-tagged GLI-1 was immunoprecipitated with an anti-haemagglutinin antibody, Myc-SUFUH (SEQ.ID.NO.12) was also present in the complex (Fig. 4a). A deletion construct containing the 407 N-terminal amino acids of GLI-1 (SEQ.ID.NO.6) also co-precipitated with SUFUH (Fig. 4b)(SEQ.ID.NO.1), whereas haemagglutinin-GLI1 did not co-precipitate with the truncated SUFUHΔ protein (SEQ.ID.NO.10)(Fig. 4c). The truncated and full-length proteins were made at similar levels (Fig. 4d). As a further control, we showed that an unrelated overexpressed protein, IκBα, does not interact with SUFUH (Fig. 4e). We conclude that the role of SUFUH as a negative regulator of GLI1-activated transcription is most likely mediated by a direct physical interaction between SUFUH and GLI-1. Our results show that the C-terminal domain of SUFUH (SEQ.ID.NO.2, SEQ.ID.NO.3) and N-terminal part of GLI-1 (SEQ.ID.NOs.6-8) are involved in protein-protein interactions, indicating that these domains may contain new types of interacting motif. The possibility that this interaction is mediated by an unidentified cellular protein appears unlikely, because all of the haemagglutinin-tagged GLI-1 could be immunoprecipitated by anti-Myc antibody from double-transfected cells, and proteins translated *in vitro* are also able to interact (data not shown). However, whether this interactions is direct or indirect remains to be confirmed.

Example 6: SUFUH can retain nuclear GLI-1 variants in the cytoplasm.

Using confocal microscopy, we determined the subcellular localization of transfected epitope-tagged GLI-1, SUFUH and SUFUHΔ in 293 cells; in each case, we studied at

least 50 positive cells, all of which showed a similar pattern. SUFUH was present both in the cytoplasm and in the nucleus (Fig. 5a). SUFUHΔ showed a very similar subcellular localization (Fig. 5b). To study GLI-1 localization, we generated an expression construct containing an N-terminal EGFP tag. Using this construct, we showed that GLI-1, like Ci, was cytoplasmic (Fig. 5c). Previous studies were not able to demonstrate the presence of full-length Ci [Aza-Blanc et al., Cell **89**:1043-1053 (1997)] in the nucleus under any conditions, a fact assumed to be caused by the presence of a cytoplasmic-retention domain. However, truncation of Ci [Aza-Blanc et al., Cell **89**:1043-1053 (1997)], GLI-1 or GLI-3 [Shin et al., Proc. Natl. Acad. Sci. USA **96**:2880-2884 (1999); Ruiz i Altaba et al., Development **126**:3205-3216 (1999)] after the DNA-binding domain leads to nuclear localization of the protein. Like these proteins, the truncated GLI1 (N407) (SEQ.ID.NO.6) construct was constitutively nuclear (Fig. 5d). This protein became partially cytoplasmic when coexpressed with SUFUH (SEQ.ID.NO.1)(Fig. 5e) but remained nuclear when coexpressed with SUFUHΔ (SEQ.ID.NO.10)(Fig. 5f).

Example 7: GLI-1 localization is regulated by nucleocytoplasmic shuttling.

We used another construct, containing the 525 N-terminal amino acids of GLI-1 (SEQ.ID.NO.15), in an attempt to map the region that mediates the cytoplasmic localization of GLI-1. This construct, like full-length GLI-1, was predominantly cytoplasmic (Fig. 6a). By inspecting the region between amino acids 407 and 525 (corresponding to the presumed cytoplasmic-retention domain) for specific sequence motifs, we found a perfect match to a leucine-rich nuclear-export signal (NES), starting at amino acid 495 (Fig. 6b). This sequence was fully conserved in other vertebrate species. In Ci we found a potential NES in the same region, as well as two extra potential NESs closer to the C terminus. The last two, but not the first, potential NESs were conserved in the mammalian Ci homologues Gli-2 and Gli-3 (Fig. 6b). To determine whether this region is indeed a functional NES, we mutated leucines 495 and 501 into alanines in this construct (SEQ.ID.NO.16) by polymerase chain reaction (PCR)-mediated mutagenesis.

The resulting protein was exclusively nuclear, indicating that this motif in GLI-1 indeed functions as an NES (Fig. 6c).

Leucine-rich NESs mediate nuclear export of proteins that is dependent on the NES-receptor protein CRM1 [Bogerd et al., J. Virol. 72:8627-8635 (1998)] which can be specifically inhibited by the drug leptomycin B (LMB) [Kudo et al., Exp. Cell. Res. 242:540-547 (1998)]. To test whether full-length GLI-1 is a target for active nuclear export, we studied the effects of LMB treatment (5 nM for 6 h) on the subcellular localization of GLI-1. Consistent with an important function for nuclear export in regulating the subcellular localization of full-length GLI-1, GLI-1 was localized to the nucleus when LMB was added (Fig. 6d). Co-transfection of SUFUH (Fig. 6e), but not SUFUHΔ (Fig. 6f), inhibited this LMB-dependent nuclear accumulation of GLI-1. It has turned out that SUFUH interacts with a subsequence from amino acid 381 to amino acid 407 of GLI-1. This interaction blocks shuttling of GLI-1.

Example 8: SUFUH can interact with GLI-1 bound to DNA.

The presence of SUFUH in the nucleus raises the possibility it might interact with GLI-1 in this cellular compartment too, potentially affecting DNA binding by GLI-1 and/or its contacts with other transcription factors. To assess this possibility, we used the constitutively nuclear GLI1(N407) variant (SEQ.ID.NO.6) in combination with SUFUH and incubated nuclear extracts from either singly or doubly transfected cells with a radiolabelled double-stranded oligonucleotide (GLI-RE) corresponding to the GLI1-binding consensus DNA sequence. We then separated the complexes by native acrylamide gel electrophoresis. GLI1(N407) specifically retards the electrophoretic mobility of the GLI-RE oligonucleotide (Fig. 7a, lane 1), as this retardation can be blocked by addition of excess unlabelled GLI-RE (lane 2). Co-transfection of SUFUH resulted in even stronger complex formation and in the appearance of a new, more retarded complex, indicating a SUFUH-GLI1 protein-protein interaction (Fig. 7a, lane 3). Antibody supershifting experiments using antibodies against the respective tags fu-

sed to GLI-1 (Fig. 7a, lane 4) and SUFUH (lane 5) sequences showed the presence of SUFUH in both retarded complexes and GLI-1 in at least the more quickly migrating complex. In support of a direct interaction between GLI-1 and SUFUH that is independent of other nuclear proteins, GLI1(N407) could be co-immunoprecipitated with SUFUH (Fig. 4b) and similar electrophoretic mobility-shift assay (EMSA) results were obtained with *in vivo*-translated proteins (data not shown). Analysis of nuclear extracts from cells transfected with SUFUH alone revealed no retarded complexes (Fig. 7a, lane 6).

The increase in DNA binding seen after co-transfection of SUFUH (SEQ.ID.NO.1) correlated with increased levels of soluble GLI1(N407) (SEQ.ID.NO.6) in the nuclear extract (Fig. 7b, lower panels). In contrast, the levels of GLI1(N407) in the nuclear pellet decreased on co-transfection with SUFUH and, as most GLI1(N407) resides in the nuclear pellet, the total nuclear level of GLI1(N407) is reduced, consistent with the immunofluorescence data (Fig. 5). GLI1(N407) protein appeared in the cytoplasm only when SUFUH was co-transfected (Fig. 7b, upper panel). Our results indicate that SUFUH can act both by preventing access of GLI-1 to the nucleus and by binding to GLI-1 when this factor is already bound to DNA.

General Methods relating to Examples 9 – 12:

Biological materials

Timely-mated NMRI mice were obtained from the Charles River Laboratories. The Dac/Dac mice and their wild-type SMNZB congenic counterparts and DNA samples from both were obtained from Jackson Laboratories. The partially overlapping BAC clones RPCI-11-170J3 containing the 5' end and 243I14 containing the 3' of the human *SU(FU)* gene were obtained from Research Genetics. The mouse *Su(fu)* BAC clones 17985 and 17986 were obtained from Genomic Systems, Inc. The mouse Su(fu)

EST clones 513730, 1195307, 1224813, and 963990 from the IMAGE consortium were obtained from the UK HGMP Resource Centre, Hinxton, Cambridge.

Chromosomal localization, mapping, and sequencing of the human *SU(FU)* and mouse *Su(fu)* genes

The chromosomal localization of *SU(FU)* was determined by Radiation Hybrid Mapping using the primer pair 5'cagttgtgtcaacgagatctcc (SEQ.ID.NO.28) and 5'ctgtggcctgtgcatggcac (SEQ.ID.NO.29) and the primer pair 5'ccgctggctaagccttgtgac (SEQ.ID.NO.30) and 5'cagggtccagagtcaaacctca (SEQ.ID.NO.31)(both primer pairs are derived from the 3' untranslated region) on the Genbridge G4 map (Research Genetics). The intron-exon borders were determined by sequencing BACs with primers oriented in both directions for each exon at least two times. The sequences of these primers are available on request. Restriction enzyme digests of BAC DNA and subcloned DNA fragments were transfered to Hybond N^+ filters (Amersham). The filters were hybridized with radiolabelled oligonucleotide probes for individual exons to assemble the Su(fu) locus map.

Screening for mRNA variants

cDNAs from normal human tissues (Human MTC panel II) or from xenografted human tumors (Human tumor MTC panel) were obtained from Clontech. Mouse embryo cDNA: Organs were homogenized and RNA was extracted using RNAzol B [Chomczynski et al., Anal. Biochem. **162**:156-159(1987)]. Subsequently, cDNA was prepared using $o(dT)_{15}$ and reverse transcriptase from mouse mammary tumor virus (Clontech) according to standard procedures. Primary PCR products were generated containing the entire coding region of SU(FU)/Su(fu) and subsequently nested PCR products containing the complete exons 2-3, 4-5, 6-7, 8-9, 10-11 respectively and flanking sequences of the neighbouring exons were obtained. After gel electrophoresis, variant transcripts were detected by size difference.

Whole mount *in situ* hybridization

Mouse embryos were cut sagitally and after fixation in 4% paraformaldehyde hybridized according to standard procedures (22)[Wilkinson, In situ Hybridization: A practical approach, IRL Press, Oxford 1992]. Digoxin-labelled probes were prepared using the DIG RNA labelling kit and the label was detected using the DIG Nucleic Acid Detection Kit (both Boehringer Mannheim). Specificity of the reaction was monitored by comparison to samples hybridized to sense probes.

Epitope-tagged constructs

Myc-SU(FU) and HA-GLI1, both in pCMV-5, have been described above. The myc-tagged constructs for the transcription variants were generated by cloning cDNAs into the pGEM-T vector (Promega), verifying the sequence of the PCR product and exchanging a fragment between two appropriate restriction sites in the myc-SUFUH. In detail, a BstEII (nt 327) -SalI (behind the STOP codon) fragment from near full length SU(FU)-XL was cloned into myc-SU(FU) and EcoNI (nt 751)-SalI fragments of SU(FU)-Tt and SU(FU)-Ly were ligated into myc-SU(FU)-XL.

Transient transfections and Western Blotting

10 µg each of the tagged GLI and SU(FU) constructs were transfected into a near confluent 15 cm culture dish with 293 kidney carcinoma cells using 75 µg of polyethyleneimine (Sigma Chemicals) in 12 ml serum-free medium as a transfecting agent. After 36 h cells were harvested and lysed in 1% Triton X-100. The lysates were precleared one time over protein A/G-agorse preincubated with goat serum and then precipitated with mouse-anti myc monoclonal antibody 9E10 or rabbit anti mouse-antiserum. After electrophoretic separation on a reducing 10% PAGE-SDS gel and blotting onto PVDF

membrane, myc- tagged proteins were detected by rabbit-anti myc antiserum, followed by incubation with a peroxidase-conjugated second antibody.

Alizarin-Alcian Blue staining

Newborn and 1 day old pups were dissected removing their skin, muscles and inner organs. Then the carcasses were stained as described [Inoue, Cong. Anom. **16**:171-173].

Sequencing of the coding regions of SU(FU) in SHFM3 patients and in the Dac mouse

mRNA was obtained from transformed patient B-cell lines or from lung of an adult homozygous Dac mouse as decribed (see above). cDNA and subsequently full length (near full length in case of Dac) PCR products of the coding regions were generated. Each two overlapping nested PCR products were generated and sequenced by standard procedures from both ends (in case of Dac the initial 49 nucleotides of the coding region are missing). At least 13 clones from each patient sample and 6 clones from the Dac mouse were included. All exons from the Dac mouse and also exon 1 from the patients were sequenced from genomic DNA using primers with intron sequences and sequences of flanking untranslated regions.

Example 9: Genomic organization of the human *SU(FU)* and of the mouse *Su(fu)* genes

The full length coding sequences of SU(FU) and Su(fu), the human and the murine homologs of the *Drosophila* signal transducer molecule Su(fu) were indentified above. By radiation hybrid mapping the human *SU(FU)* gene was localized to chromosome 10q24 in a region syntenic to chromosome 19 in the mouse. We identified artificial bacterial chromosomes (BACs) containing both the human *SU(FU)* and the murine

Su(fu) genes and determined the intron-exon organization by sequencing using primers from the coding region.

Both genes contain twelve exons with exon-intron boundaries at corresponding positions (Table 1). When mapping the overlapping BAC clones 17985 and 17986, the mouse *Su(fu)* gene was found to span approximately 100 kb (Fig. 9). The promoter region contained no consensus TATA- or CAAT-box immediately upstream of the putative transcriptional start site. However, using the Transcription Start Site Wingender database (TSSW) at UK HGMP Resource Centre (www.hgmp.mrc.ac.uk) a promoter in this region with an excellent prediction value (LDF=20.80) was found. When sequencing upstream both the human and murine promoters, the ATG-containing exon for the gene *actin-related protein 1* (*ARP1* and *Arp1* in human and mouse respectively) was detected in opposite transcriptional orientation separated by only 1516 bp (in human) between the initial methionine codons. The ubiquitously expressed ARP1 [Clark et al., Mol. Biol. Cell **5**:1301-1310 (1994)] and Arp1 [Schroer et al., J. Cell Biol. **115**:1309-1319 (1991)] are members of the microtubule-associated dynactin complex which is essential for spindle formation [Merdes et al., J. Cell Biol. **149**:851-862 (2000)] and organelle transport [Xiang et al., Curr. Biol. **18**:603-606 (2000)], for a review see [Schafer et al., Annu. Rev. Cell. Dev. Biol. **15**:341-363 (1999)] (Fig. 9). In addition, sequencing off the end of BAC 17986 revealed an exon for the *ADP-ribosylation factor-like protein 3* (*Arl3*) downstream of the mouse *Su(fu)* gene (Fig. 9). Mining the GenBank mouse EST database, several Su(fu) EST clones originating from different tissues including adult testis, heart, kidney, B cell, mammary gland, day post natum (dpn) 6 head, dpn 10 skin, fertilized egg, embryos from dpc 8, 10-11, and 13 were found, demonstrating expression in a wide range of tissues. Four mouse ESTs were sequenced and their structure is shown in Fig. 9B. The combined EST cDNAs extend over 4400 bp, which fits well with the mRNA size seen on Northern blot (4.5 kb). The predicted AUG codon of the translational initiation has a good Kozak consensus sequence. The clones 1224813 and 963990 that extends the most 3' as well as a number of other human and mouse ESTs in the GenBank end at the same position. In-

stead of the consensus polyadenylation signal AAUAAA we found an AUUAAA sequence about 15 bp upstream of the end of the longest ESTs in both human and mouse.

Example 10: SU(FU) mRNA isoforms

We screened a commercially available panel of cDNAs from various tissues of apparently healthy donors for expression of variants of SU(FU). One variation which we found was a trinucleotide insertion after nucleotide 756 of the coding sequence resulting in an extra-glutamine after amino acid 252. This insertion created a novel EcoNI cleavage site which allowed easy detection of this variant which we called SU(FU)-XL (SEQ.ID.NO.17). SU(FU)-XL was found to be expressed in five of seven tissues tested with a relative abundance varying from roughly 10% to 50% of total SU(FU) (Fig. 10). The explanation for this variation was found in the sequence of the downstream end of intron 6 which was determined as ...TTTTCAAGCAG-exon 7... (Table 1). Probably, both underlined AG dinucleotides could function as alternative splice acceptor sites.

We found two additional splice variants (Fig. 10), one of them, SU(FU)-LK (SEQ.ID.NO.18), was expressed in peripheral leukocytes, the other one, SU(FU)-TT (SEQ.ID.NO.19), in testes. By sequencing it was found that SU(FU)-LK is lacking exon 10, and that the downstream sequence has a shifted reading frame resulting in a STOP codon after two amino acids. SU(FU)-TT has an extra exon after exon 8 (exon 8a), which encodes 19 amino acids followed by a STOP codon.

We also screened a commercially available panel of tumor cell xenograft cDNAs (two breast, two lung, two colon, one prostate and one pancreatic carcinoma sample) and cDNA samples from a panel of dpc 13.5 mouse embryonic tissues (lung, intestine, skeletal muscle, heart, liver, eye, brain, spinal cord, tongue, kidney, whole limb, skin) and we detected the normal SU(FU)/Su(fu) transcripts but no variants except SU(FU)-XL. There may be additional isoforms, since our method did not allow detection of isoforms, where the first or the last exon of SU(FU) is missing.

Functional analysis of these SU(FU) variants compared with the standard SU(FU) (as we call the originally cloned variant) was performed by assessing their potential to co-precipitate GLI-1, one of the physiological binding partners of SU(FU). N-terminally myc-tagged versions of the coding sequences of all SU(FU) variants and a hemaglutinin-tagged version of GLI-1 (HA-GLI) were transfected into 293 cells. With an antibody against HA we could coprecipitate SU(FU) when cotransfected with myc-SU(FU) or myc-SU(FU)-XL, but not with myc-SU(FU)-LK or myc-SU(FU)-TT as detected by an anti-myc antibody after Western Blotting (Fig. 11).

Example 11: Expression of SU(FU) during mouse embryogenesis

We analyzed mouse embryos in the time range between dpc 8.5 and dpc 15.5 for expression of SU(FU) by whole mount in situ hybridization. Two other reports have described similar experiments [Stone et al., J. Cell Science **112**:4437-4448 (1999); Pearse et al., Developmental Biology **212**:323-336 (1999)]. Here we will focus on those results not described before.

Above, SU(FU) has been shown to be expressed ubiquitously in the adult. A similar result has been obtained for the mouse embryo [Stone et al., J. Cell Science **112**:4437-4448 (1999); Pearse et al., Developmental Biology **212**:323-336 (1999)] and our studies confirm these results in principle. In particular at dpc 10.5 and 11.5 we observed an intense all over reactivity of which only the neural tube and the forming CNS emerged as even stronger and only the liver appeared to be negative. For the other stages however, we found that certain tissues expressed Su(fu) with considerably stronger intensity than the all over reactivity of the embryos. At dpc 8.5 and 9.5 we detected Su(fu) in the neural tube and in the somites (Fig. 12A-C). At dpc 9.5 Su(fu) was expressed further in the heart and the dorsal artery as well as in the intersegmentary arteries (Fig. 12C, F), in the bud of the forelimb (Fig. 12B, C), in the branchial arches (Fig. 12B, F), and in the mesenchyme of the trunk and of the head surrounding the prospective mouth cavity (Fig. 12B). Su(fu) expres-

sion in the neural tissues persisted through the entire observed period as previously reported [Stone et al., J. Cell Science **112**:4437-4448 (1999); Pearse et al., Developmental Biology **212**:323-336 (1999)]. At dpc 10.5 Su(fu) was expressed in all brain vesicles (Fig. 12D) and in the dorsal root ganglia (Fig. 12E) as detected after removal of the skin. At all later stages, the brain and the spinal cord were stained with high intensity (Fig. 12G). The expression of Su(fu) in the somites persisted as well, however from dpc 10.5 only the sclerotome and the sclerotome-derived tissues were stained with the probe. The mesenchyme surrounding the ossifying part of the vertebrae expressed Su(fu) from dpc 12.5 (Fig. 12G) and of the ribs from dpc 13.5 [Pearse et al., Developmental Biology **212**:323-336 (1999)]. Like in the somite-derived skeletal parts, even in the digits the expression of Su(fu) was observed to parallel the progress of ossification as the area which expresses Su(fu) was migrating distally during the course of development (Fig. 12H-J). The expression of Su(fu) in the branchial arches was continuing within the maxilla, mandible and in the tongue and disappearing from dpc 12.5 onwards when the mouth cavity was closed, with at dpc 15.5 only tongue being stained (Fig. 12K). Further, through the entire observed period Su(fu) was expressed in the genital tubercle (Fig. 12G) and from dpc 13.5 in the mesenchyme surrounding the developing tracheoles of the lung (Fig. 12L).

Example 12: Examination of RNA from SHFM3 patients and DNA from Dac mice for mutations in the *Su(fu)* gene

We have located the *SU(FU)* gene in man between the markers AFM183XB12 and D10S192, with a distance of 0.10 cR from AFM183XB12 (lod>3.0). This narrow region also contains the locus for the genetic disorder Split Hand Split Foot-Malformation-3 (SHFM3). Members of the Hedgehog signalling pathway have been shown to be involved in the process of digit formation as will be discussed below.

To investigate the hypothesis that *SU(FU)* would be the gene involved in SHFM3, we sequenced a number of cDNA clones of SU(FU) from B-cell lines derived from two unrelated heterozygous SHFM3 patients (since we met difficulties in reverse transcrip-

tion of the first exon we sequenced the first exon from the genomic DNA of these cell lines.). However, in the entire coding sequence no differences from the normal sequence were observed. With our approach mutations in the promotor or in introns which affect the transcriptional regulation or splicing of *SU(FU)* can not be excluded.

In the mouse, the Dac mutant [Chai, J. Hered. 72:234-237 (1981)] with a similar phenotype [Crackover et al., Dev. Biol. **201**:78-89 (1998); Seto et al., Teratology **56**:161-270 (1997)], is linked to chromosome 19 in a region syntenic to the one on human chromosome 10 containing the locus for the *SHFM3* gene. It has been suggested that Dac might be the murine counterpart of SHFM3 [Johnson et al., Genomics **29**:457-464 (1995)]. We amplified and sequenced all Su(fu) exons from the genomic DNA of a homozygous Dac mouse. Additionally, the coding region of the cDNA from nucleotide 50 after the initial AUG codon downstream was sequenced. No mutations in the *Su(fu)* gene were observed and the presence of a normal mRNA sequence makes splice mutations very unlikely.

During our studies on Dac mice we observed abnormalities not only in the digit formation but also in cranial development. The ossification of the skull bones of the Dac mouse in the perinatal period was delayed compared to the wildtype. The intramembraneously developing parietal, interparietal and supraoccipital bones have not been ossified at all even at dpn 1 (two mice examined at dpn 0, one at dpn 1), while they have been ossified completely at dpn 0 in the wild-type control mice. The endochondrally developing frontal, exoccipital and temporal bones, as well as the basal processes and the parts of the jaw, however, were developed in the Dac mouse as well as in wild-type (Fig. 13). Further, in the newborn Dac/Dac mouse the skull was flat and less curved than in wild-type.

Table 1. Exon-intron boundaries of the human *SU(FU)* and mouse *Su(fu)* genes

Exon number	Exon	Intron (length)
1	1-ATGGCG...GTACTG-182	gtatgc...ttgcag
	1-ATGGCG...GTACTG-182	gtctgc...ttgcag (3.7 kb)†
2	183-GTTGGG...CCATGA-317	gtgaag...tttcag
	183-GTTGGG...CCATGA-317	gtgagt...tttcag (24 kb)†
3	318-GTTTAC...AGTCAG-454	gtagga...ccacag
	318-GTTTAC...AGTCAG-454	gtaata...ccacag (22 kb)†
4	455-AGAACA...CTCCAG-597	gtgagg...cctcag
	455-AGAACA...CTCCAG-597	gtgagg...ccacag (607 bp)§
5	598-ATCGTT...GCCTAT-683	gtgagt...ccacag
	598-ATTGTT...GCCCAT-683	gtgagt...ccacag (226 bp)§
6	684-TGCTGG...CTGCAA-756	gtatgt...aagcag™
	684-TGCTGG...CTGCAA-756	gtatgt...aagcag™ (2.3 kb)†
7	757-GAGAGA...GCAAAG-910	gtggga...ttgcag
	757-GAGAGA...GCAAAG-910	gttgag...ttgcag (1419 bp)§
8	911-ACACAG...GGCCCC-1022	gtaagt...tcacag
	911-ACACAG...GGCTCC-1022	gtaagt...tcacag (22 kb)†
9	1023-GAGCCG...CCTAAG-1157	gtgagc...ctccag
	1023-GAGCCG...CCTAAG-1157	gtgagc...ccccag (1.6 kb)†
10	1158-GGGCAG...TTACAA-1296	gtgaga...tcacag
	1158-GGGCAG...TTACAA-1296	gtgaga...tcccag (9.9 kb)†
11	1297-ATTCTGTTGA...GAGGAA-1365	gtaagc...ccacag¶
	1297-ATTCTG...GAGGAA-1365	gtgagc...tcacag (1.9 kb)†
12	1366-TTCAAA...	
	1366-TTTAAA...CTGCAG-4309	

Note: The upper and lower row in each pair refers to the human and mouse sequence, respectively. The translational start site is numbered as nucleotide 1 since the transcriptional start site is not exactly defined. Exon 1 contains approximately an additional 150 bp 5´of the translational start site.

§Fully sequenced.

†Partially sequenced.

™The dotted underlined ag is the intron border in SU(FU)-XL

¶The double-underlined TGA is the STOP codon in SU(FU)-LK

Claims

1. A peptide consisting of one or more multiples of a protein fragment comprising 5 – 104 amino acid residues, preferably 10 – 50 amino acid residues, derived from the amino acid sequence of SEQ.ID.NO.2, said protein fragment showing a homology of at least 70 %, preferably at least 90 %, with the amino acid sequence of SEQ.ID.NO.2, said peptide being able to spercifically bind to a protein having the sequence of SEQ.ID.NO.5.

2. A peptide according to claim 1, **characterised** in that it consists of one or more multiples of a protein fragment consisting of 10 – 97 consecutive amino acid residues from the amino acid sequence of SEQ.ID.NO.3.

3. A peptide according to claim 2, **characterised** in that said protein fragment has the amino acid sequence of SEQ.ID.NO.4.

4. A peptide consisting of one or more multiples of a protein fragment comprising 5 – 407 amino acid residues, preferably 5 – 50 amino acid residues, derived from the amino acid sequence of SEQ.ID.NO.6, said protein fragment showing a homology of at least 70%, preferably at least 90%, with the amino acid sequence of SEQ.ID.NO.6, said peptide being able to specifically bind to a protein having the sequence of SEQ.ID.NO.1.

5. A peptide according to claim 4, **characterised** in that it consists of one or more multiples of a protein fragment consisting of 5 – 26 consecutive amino acid residues from the amino acid sequence of SEQ.ID.NO.7, and which fragment contains the amino acid subsequence lys-arg-his-arg.

6. A peptide according to claim 4, **characterised** in that it consists of one or more multiples of a protein fragment consisting of 5 – 20 consecutive amino acid residues from the amino acid sequence of SEQ.ID.NO.8.

7. A DNA sequence encoding a peptide according to anyone of claims 1 – 6.

8. A monoclonal antibody or an antibody fragment directed against a peptide according to anyone of claims 1 – 6.

9. A peptide according to anyone of claims 1 – 6, for medical use.

10. Use of a peptide according to anyone of claims 1 – 6, for preparing a pharmaceutical composition for treating cancer.

11. A monoclonal antibody or an antibody fragment according to claim 8, for medical use.

12. Use of a monoclonal antibody according to claim 8, for preparing a pharmaceutical composition for treating cancer.

13. A pharmaceutical composition comprising a peptide according to anyone of claims 1 – 6 and/or a monoclonal antibody or an antibody fragment according to claim 8, together with a pharmaceutically acceptable carrier, excipient or diluent.

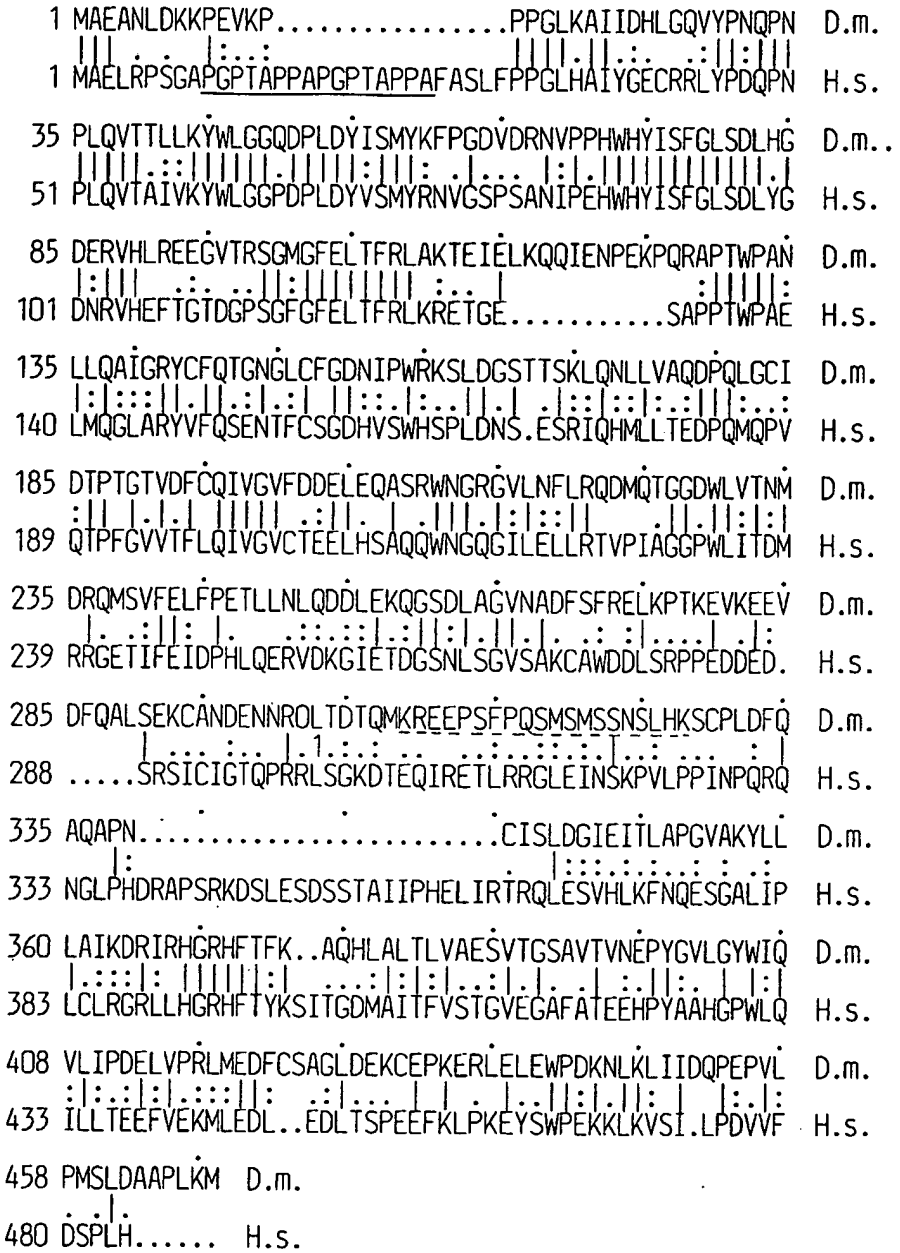

1 MAEANLDKKPEVKP.....PPGLKAIIDHLGQVYPNQPN D.m.
1 MAELRPSGAPGPTAPPAPGPTAPPAFASLFPPGLHAIYGECRRLYPDQPN H.s.
35 PLQVTTLLKYWLGGQDPLDYISMYKFPGDVDRNVPPHWHYISFGLSDLHG D.m..
51 PLQVTAIVKYWLGGPDPLDYVSMYRNVGSPSANIPEHWHYISFGLSDLYG H.s.
85 DERVHLREEGVTRSGMGFELTFRLAKTEIELKQQIENPEKPQRAPTWPAN D.m.
101 DNRVHEFTGTDGPSGFGFELTFRLKRETGE.....SAPPTWPAE H.s.
135 LLQAIGRYCFQTGNGLCFGDNIPWRKSLDGSTTSKLQNLLVAQDPQLGCI D.m.
140 LMQGLARYVFQSENTFCSGDHVSWHSPLDNS.ESRIQHMLLTEDPQMQPV H.s.
185 DTPTGTVDFCQIVGVFDDELEQASRWNGRGVLNFLRQDMQTGGDWLVTNM D.m.
189 QTPFGVVTFLQIVGVCTEELHSAQQWNGQGILELLRTVPIAGGPWLITDM H.s.
235 DRQMSVFELFPETLLNLQDDLEKQGSDLAGVNADFSFRELKPTKEVKEEV D.m.
239 RRGETIFEIDPHLQERVDKGIETDGSNLSGVSAKCAWDDLSRPPEDDED. H.s.
285 DFQALSEKCANDENNRQLTDTQMKREEPSFPQSMSMSSNSLHKSCPLDFQ D.m.
288 .....SRSICIGTQPRRLSGKDTEQIRETLRRGLEINSKPVLPPINPQRQ H.s.
335 AQAPN.....CISLDGIEITLAPGVAKYLL D.m.
333 NGLPHDRAPSRKDSLESDSSTAIIPHELIRTRQLESVHLKFNQESGALIP H.s.
360 LAIKDRIRHGRHFTFK..AQHLALTLVAESVTGSAVTVNEPYGVLGYWIQ D.m.
383 LCLRGRLLHGRHFTYKSITGDMAITFVSTGVEGAFATEEHPYAAHGPWLQ H.s.
408 VLIPDELVPRLMEDFCSAGLDEKCEPKERLELEWPDKNLKLIIDQPEPVL D.m.
433 ILLTEEFVEKMLEDL..EDLTSPEEFKLPKEYSWPEKKLKVSI.LPDVVF H.s.
458 PMSLDAAPLKM D.m.
480 DSPLH..... H.s.

FIG. 1a

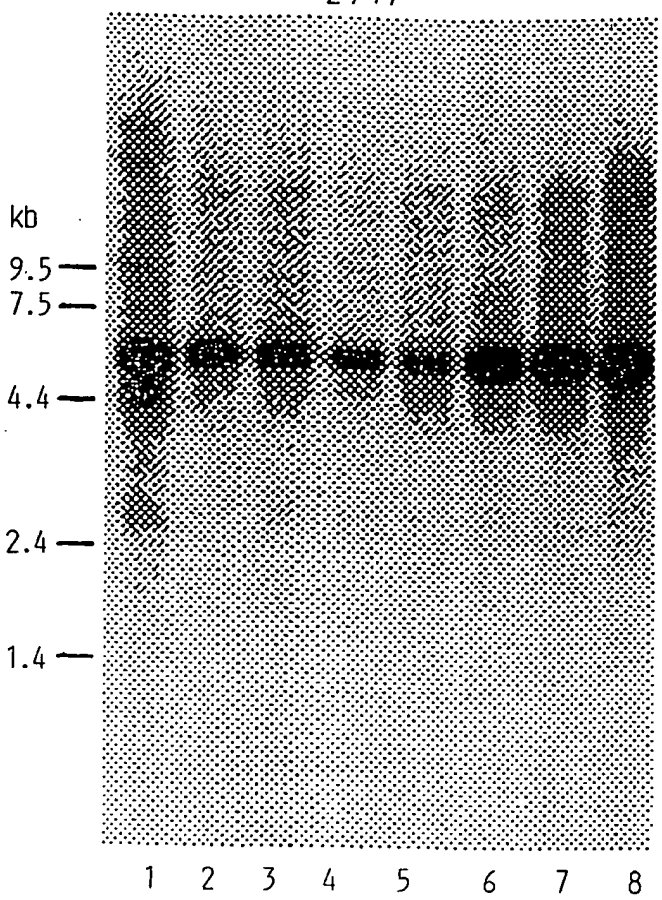
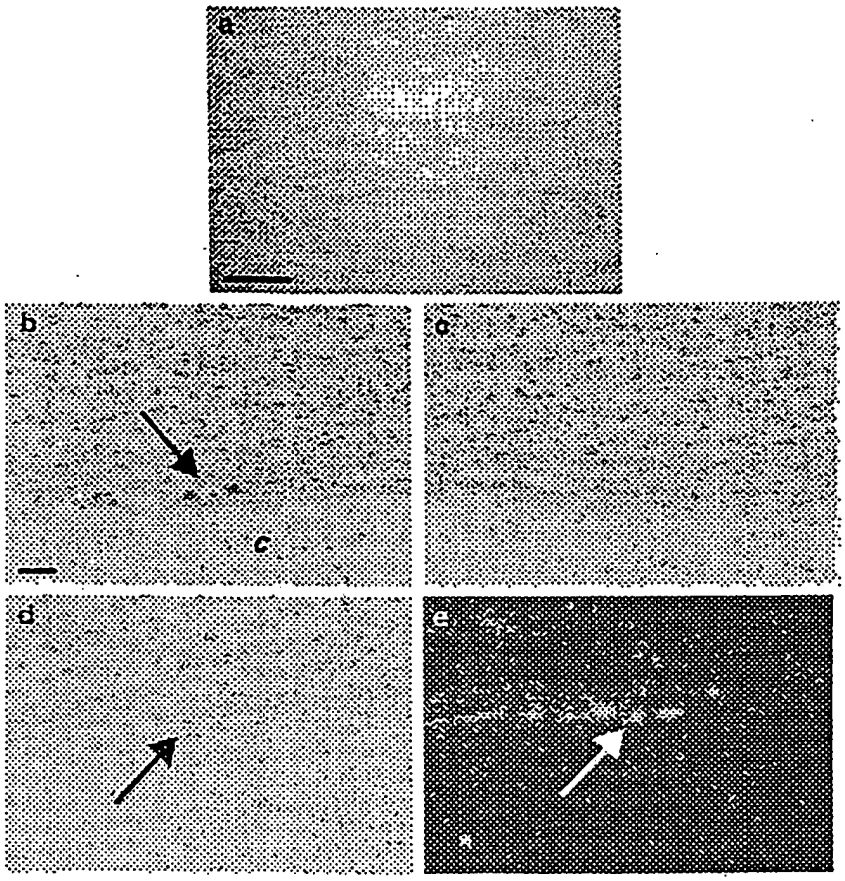


FIG.1b

FIG.1c

Fig. 2



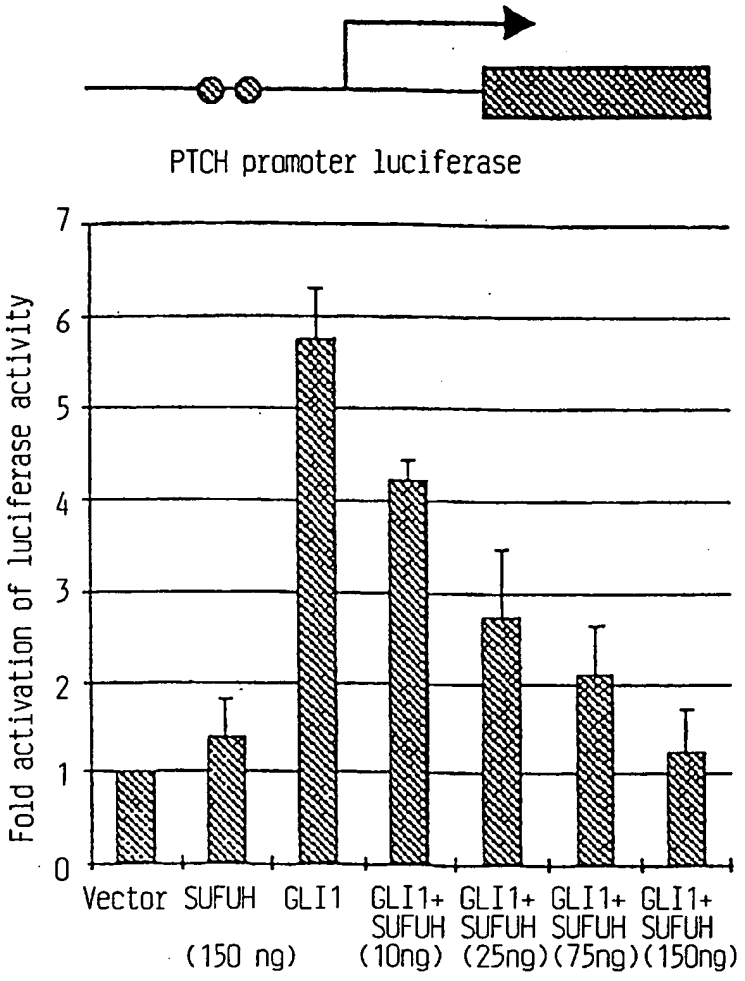
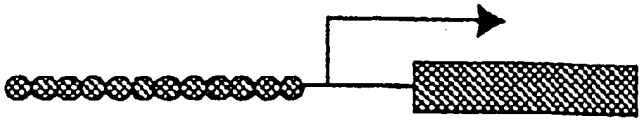

PTCH promoter luciferase
Fold activation of luciferase activity
0
1
2
3
4
5
6
7
Vector
SUFUH (150 ng)
GLI1
GLI1+ SUFUH (10ng)
GLI1+ SUFUH (25ng)
GLI1+ SUFUH (75ng)
GLI1+ SUFUH (150ng)

FIG. 3a



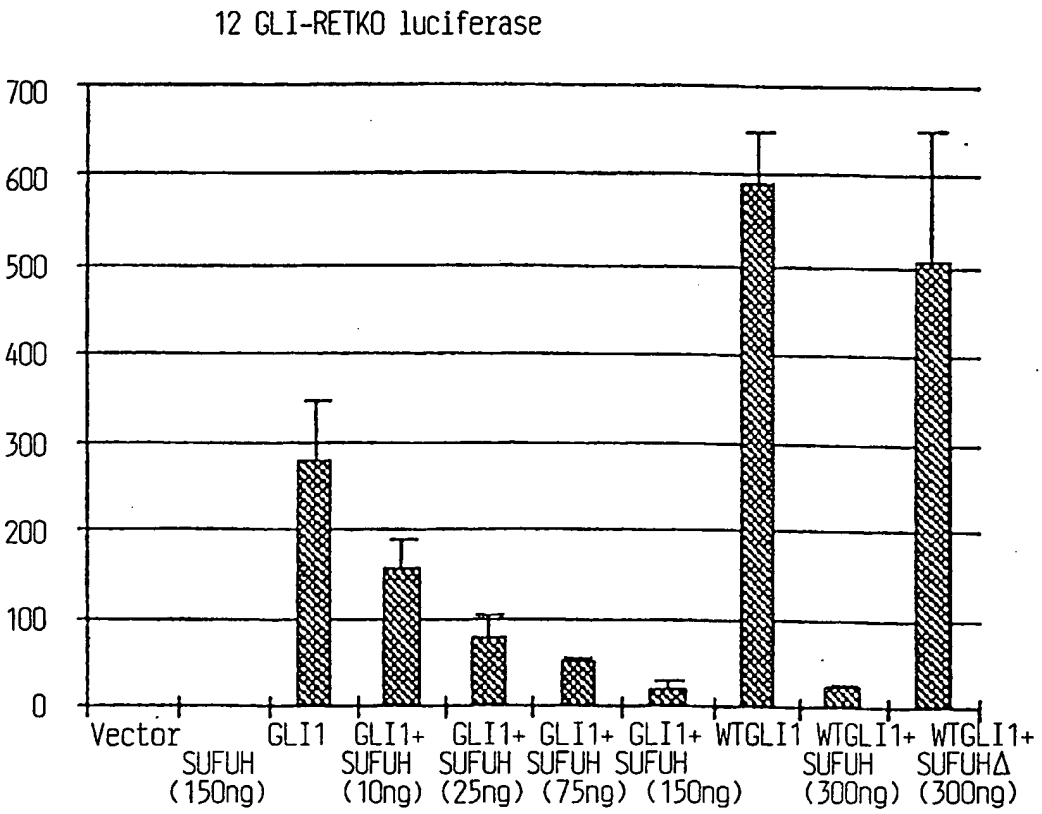
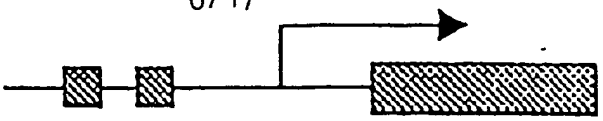

12 GLI-RETKO luciferase
700
600
500
400
300
200
100
0
Vector
SUFUH (150ng)
GLI1
GLI1+ SUFUH (10ng)
GLI1+ SUFUH (25ng)
GLI1+ SUFUH (75ng)
GLI1+ SUFUH (150ng)
WTGLI1
WTGLI1+ SUFUH (300ng)
WTGLI1+ SUFUHΔ (300ng)

FIG.3b



2 Gluc-RETKO luciferase

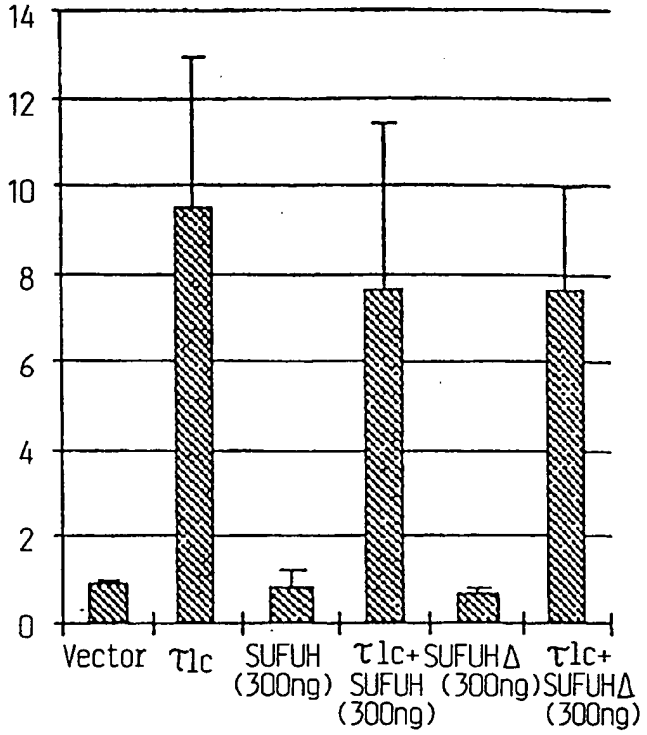


FIG.3c

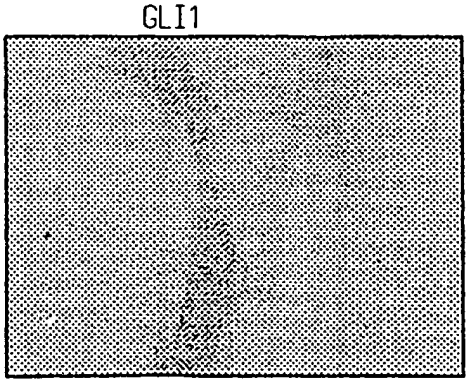


FIG. 3d

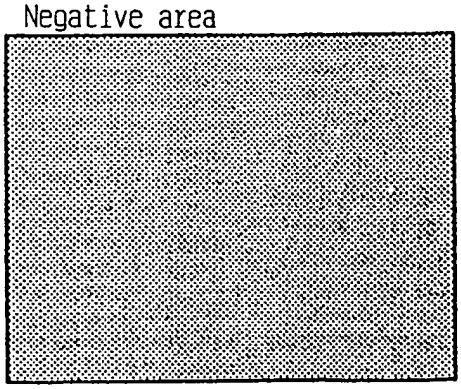


FIG.3e

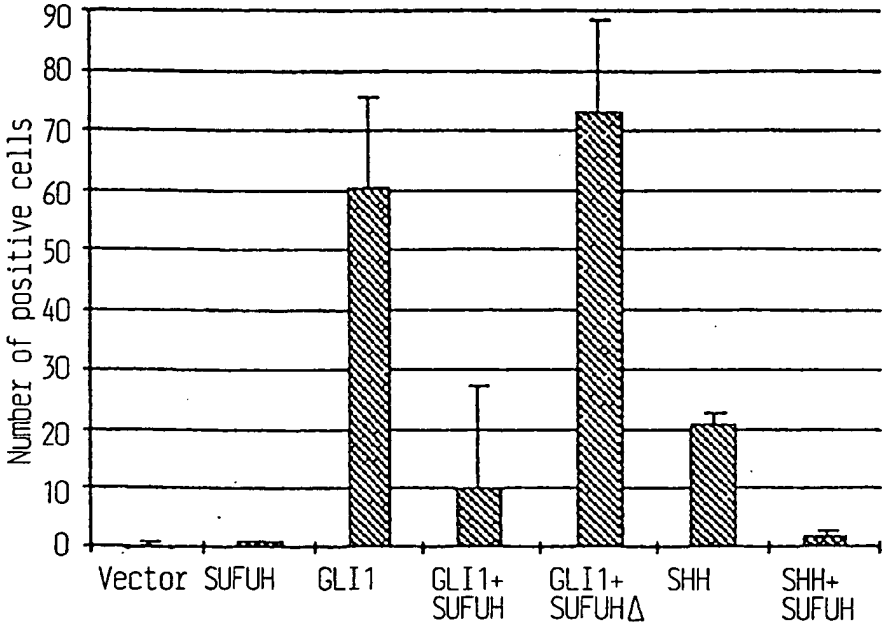

90
80
70
60
50
40
30
20
10
0
Number of positive cells
Vector
SUFUH
GLI1
GLI1+ SUFUH
GLI1+ SUFUHΔ
SHH
SHH+ SUFUH

FIG.3f

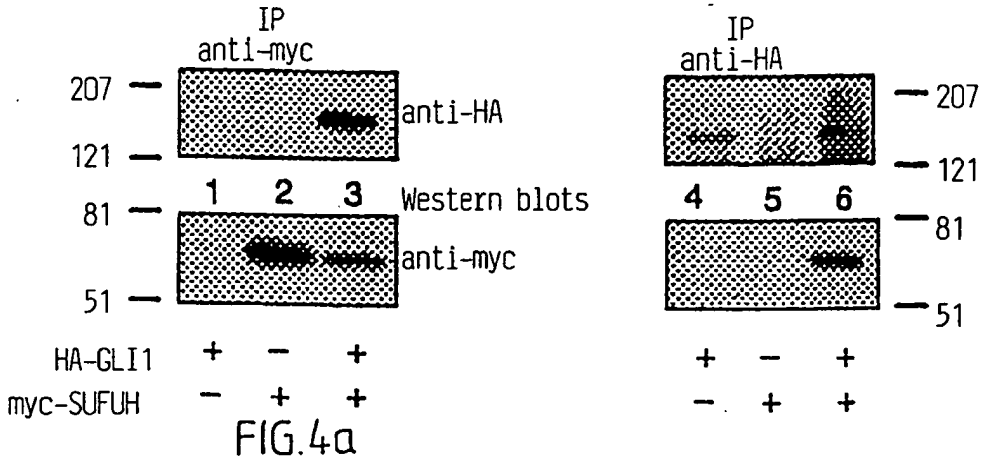


FIG.4a

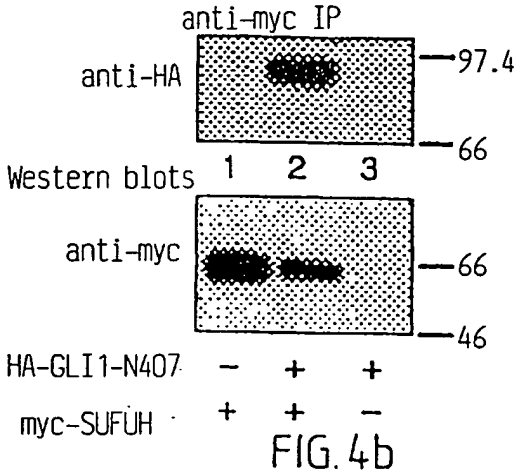


FIG.4b

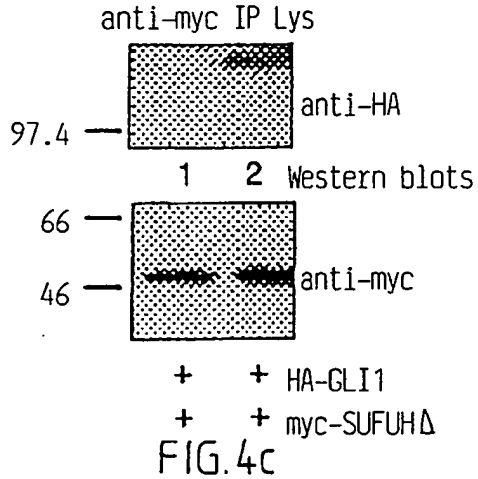


FIG.4c

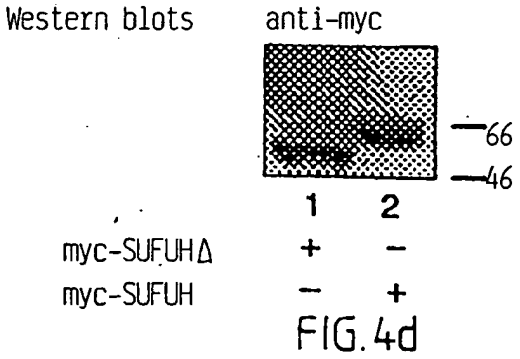


FIG.4d

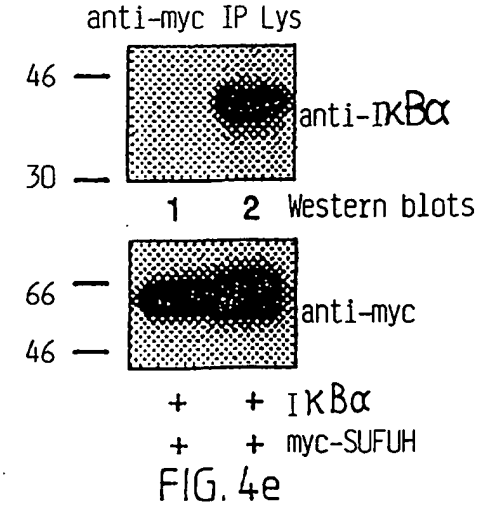


FIG.4e

Fig. 5

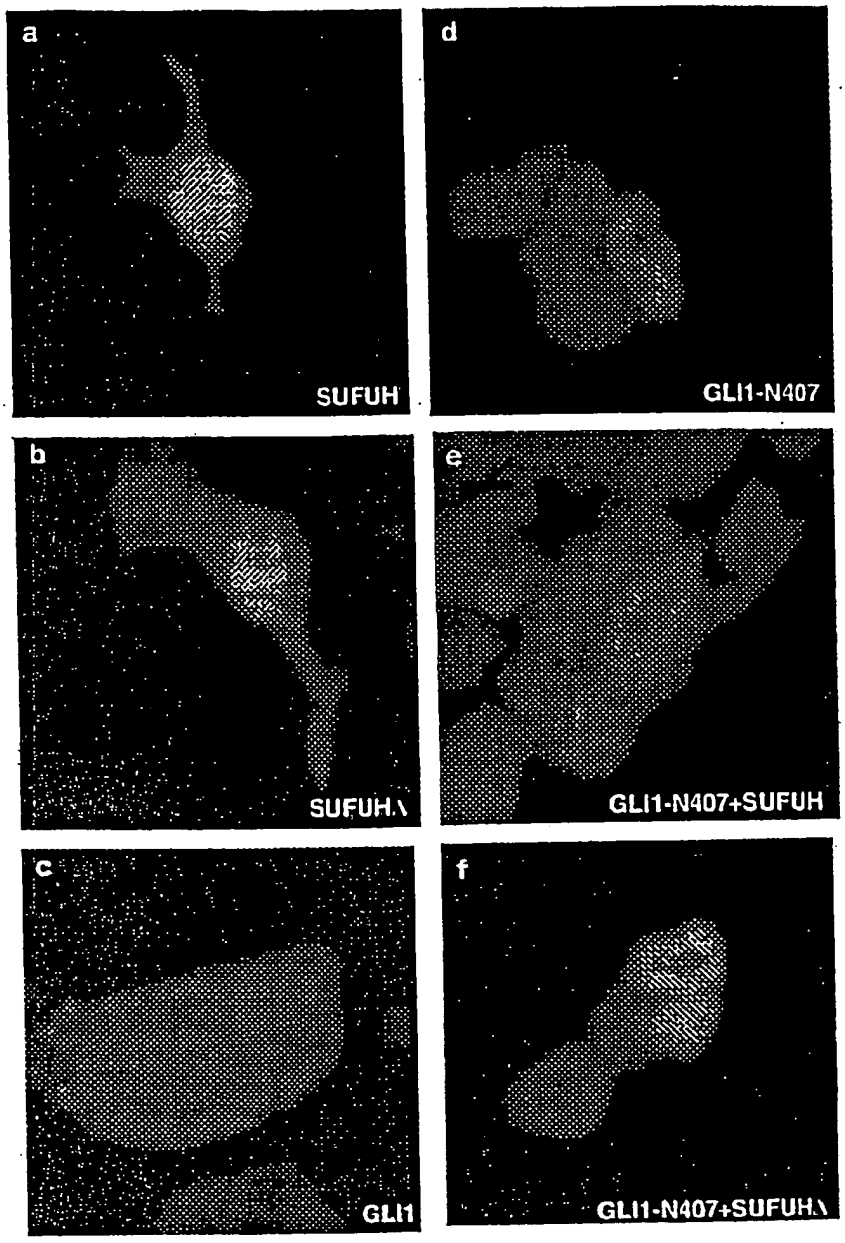
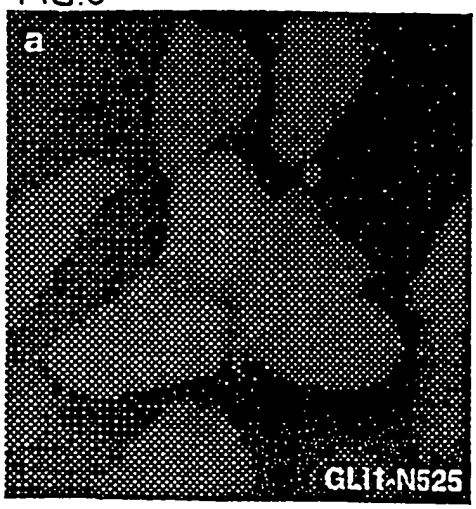
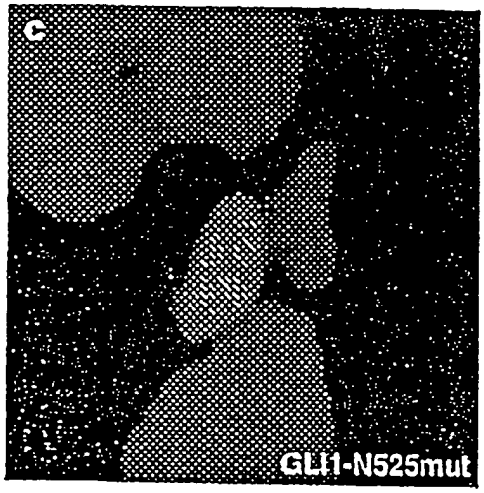

a
SUFUH
d
GLI1-N407
b
SUFUH.\
e
GLI1-N407+SUFUH
c
GLI1
f
GLI1-N407+SUFUH.\

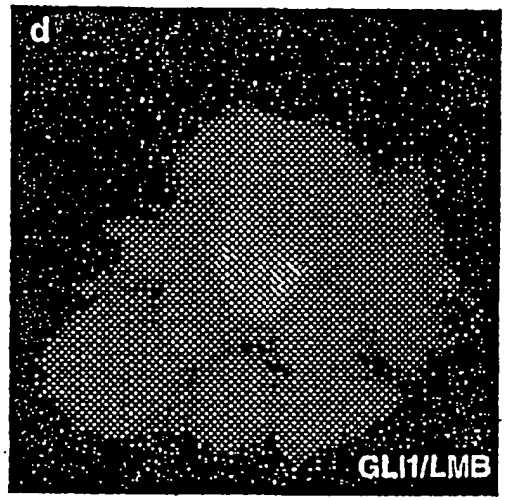
FIG.6

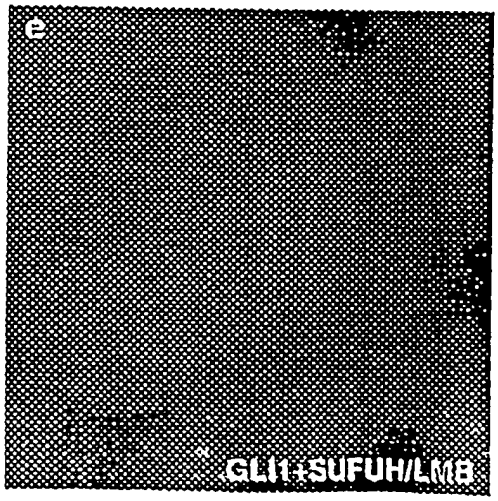


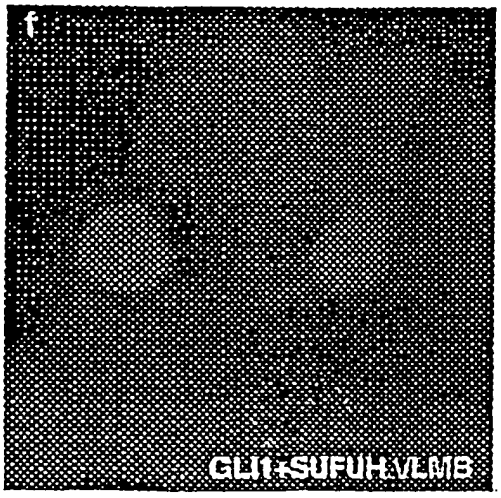
b

GLI1	Human	495	LRRLENLRL
	Mouse	496	LRRLENLRL
	Xenopus	517	LRKLENLRI
	Chicken	512	LQKLENLRI
Ci (Drosophila)		774	LNQRI.TELKM
		1,049	LRKDL.FTVSI
		1,084	LPDEMLQYLNL
GLI2		577	LPG.LER.MSL
GLI3		945	LPN.MER.MSL
		960	LGDALEPGVAL









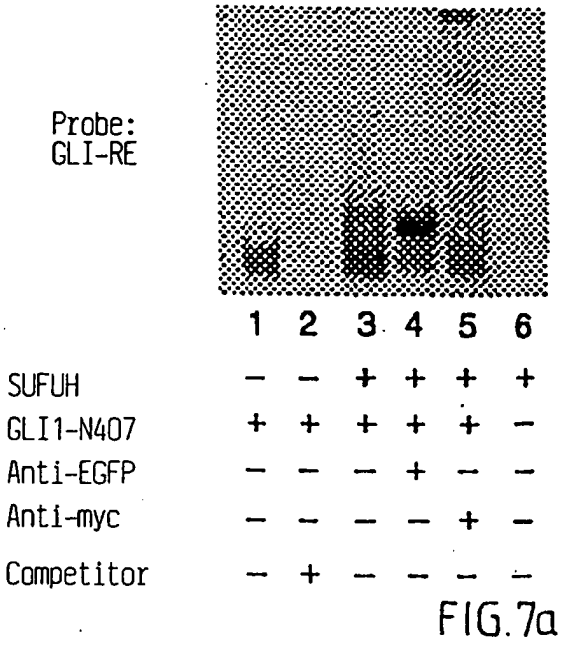

Probe:
GLI-RE
1 2 3 4 5 6
SUFUH − − + + + +
GLI1-N407 + + + + + −
Anti-EGFP − − − + − −
Anti-myc − − − − + −
Competitor − + − − − −

FIG.7a

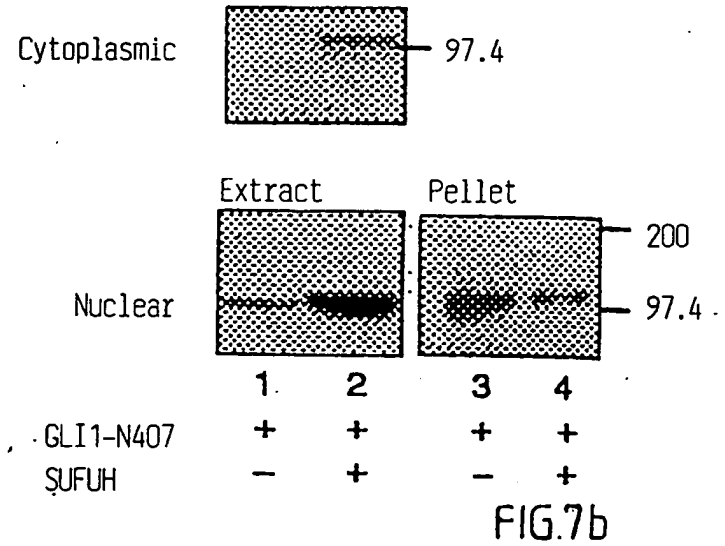

Cytoplasmic
97.4
Extract
Pellet
200
Nuclear
97.4
1 2 3 4
GLI1-N407 + + + +
SUFUH − + − +

FIG.7b

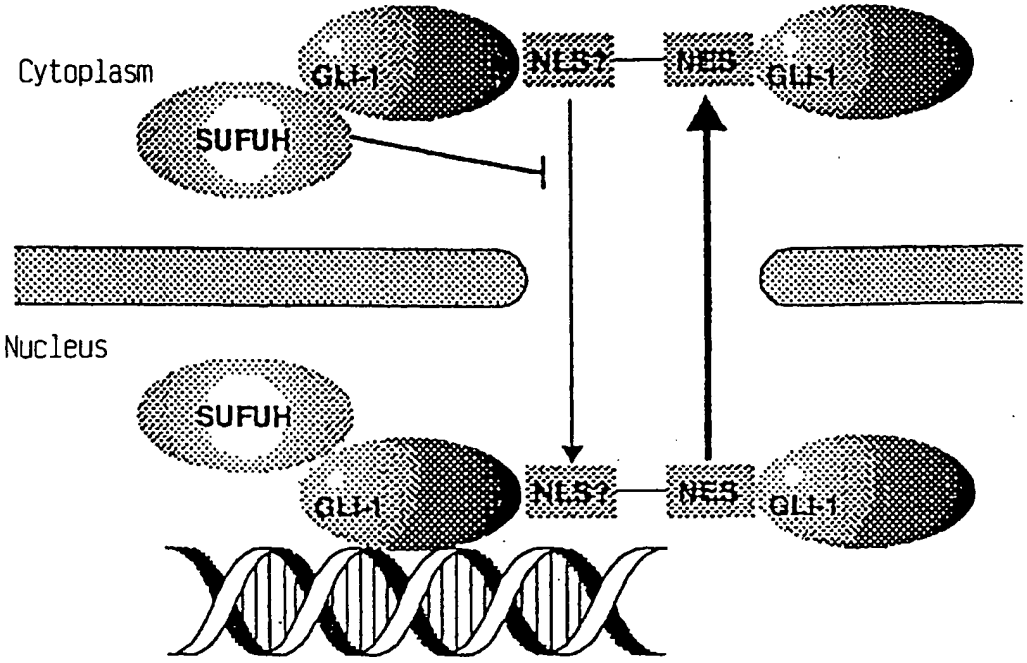

Cytoplasm
GLI-1
NLS?
NES
GLI-1
SUFUH
Nucleus
SUFUH
GLI-1
NLS?
NES
GLI-1

FIG. 8

Fig. 9

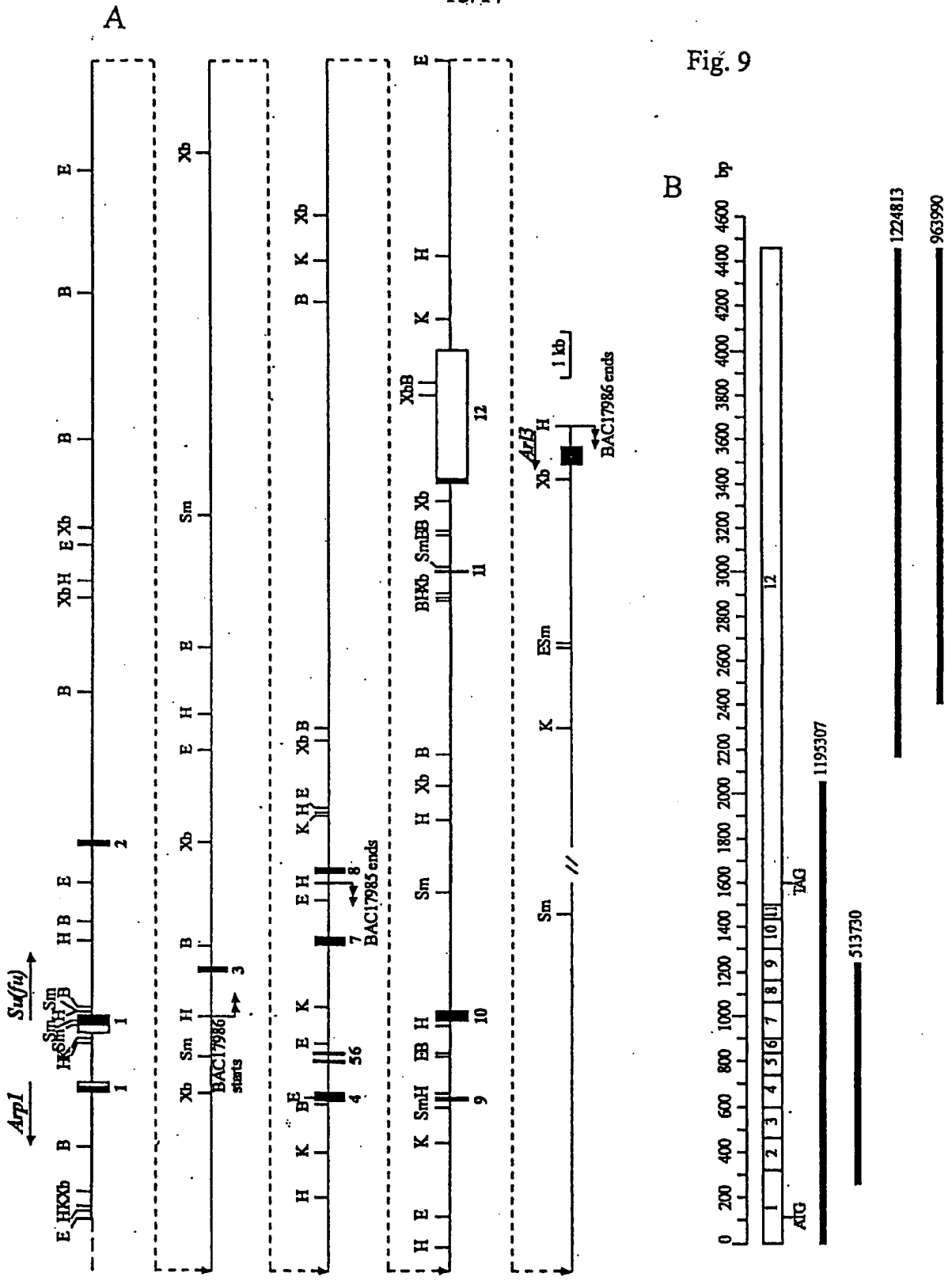
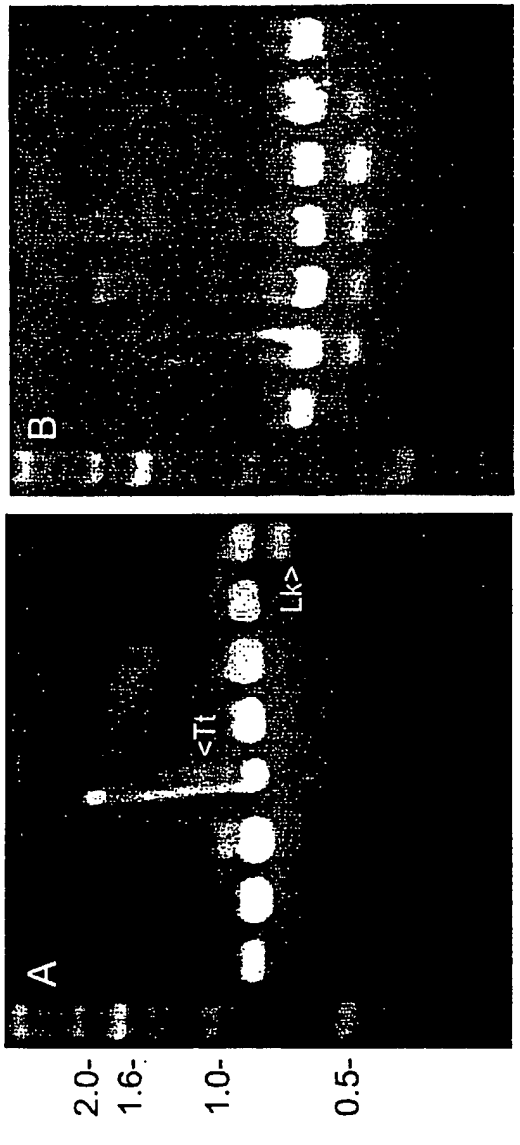

A
Arp1
Suf(u)
BAC17986 starts
BAC17985 ends
BAC17986 ends
1 kb
B
bp
ATG
TAG
1195307
513730
1224813
963990

Fig. 10



C

GCCCCAGATGAAGTCACGGCCATTGTTAACACCTCAACTTCAGAATTCCTTGCTGGTGCTTGAAATT
A P D E V T A I V N N T S S T S E F L A G A STOP
CTCAGTCCAAACGCTGCATCCCCCTCCCTCGAGTTGCACAATCAGGGGTATTTTTATCGCCCATCCT
TGATCCTCGGAGTCACCCAACACTTCAGCAGTGAGTCACTTCTCCAGGGAAGGGAAAAGGAACGCTC
AAAGGAGCCC

Fig. 11

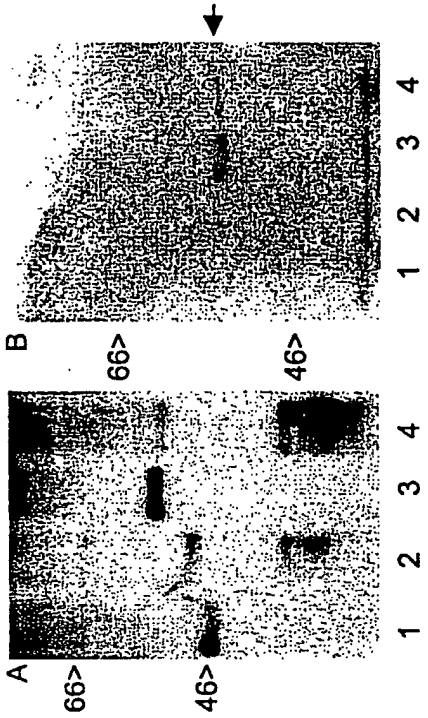


Fig. 12

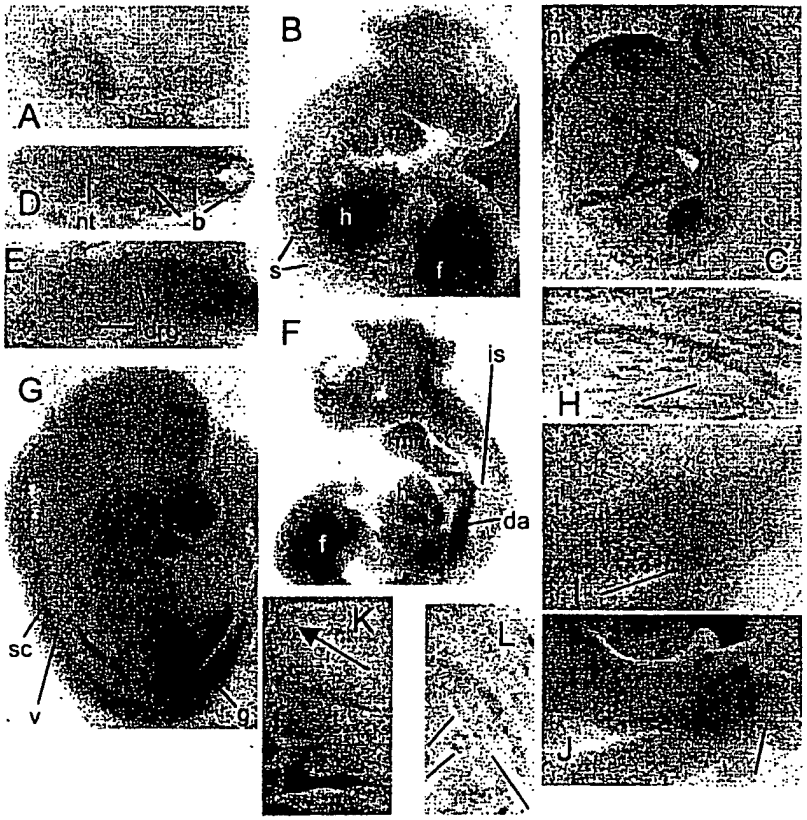
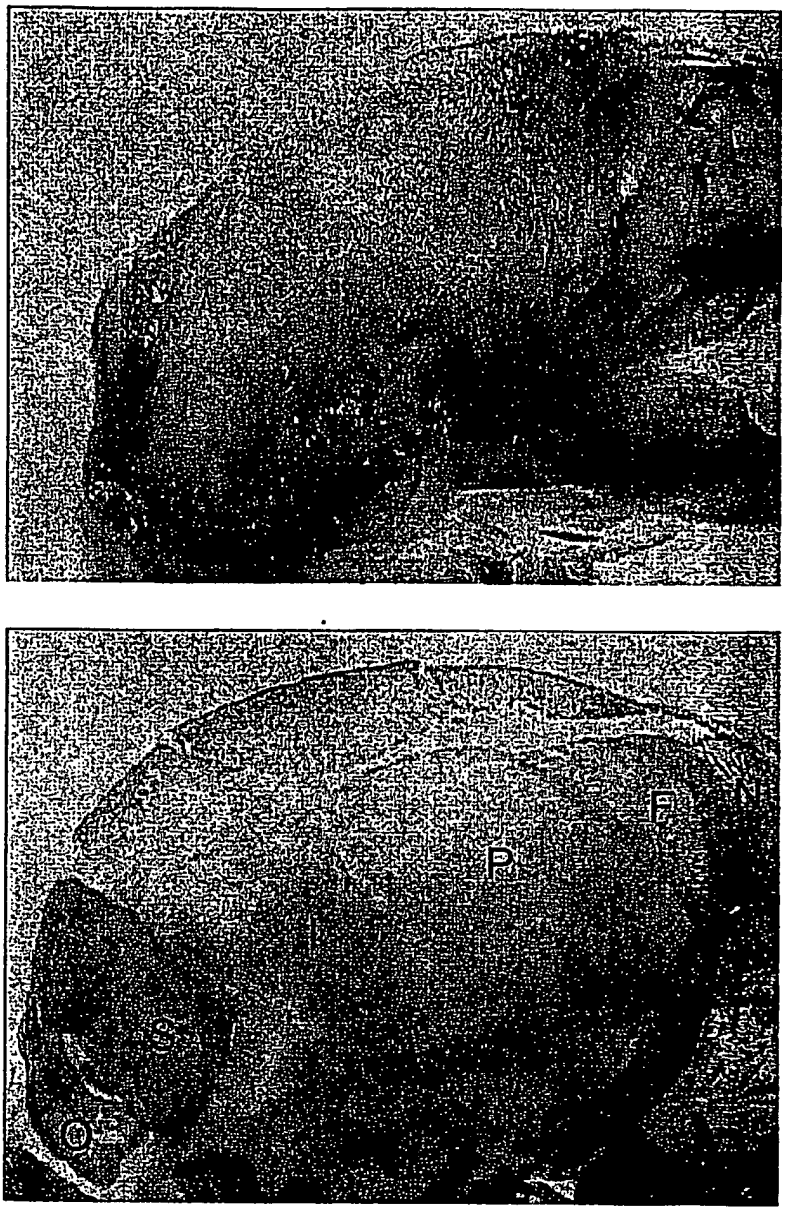


Fig. 13



SEQUENCE LISTING

<110> Karolinska Innovations AB

<120>

<130>

<150> SE 9902899-S
<151> 1999-08-13

<160> 31

<170> PatentIn Ver. 2.1

<210> 1
<211> 484
<212> PRT
<213> Homo sapiens

<400> 1

```
Met Ala Glu Leu Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro
  1               5                  10                  15

Ala Pro Gly Pro Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro
             20                  25                  30

Gly Leu His Ala Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln
         35                  40                  45

Pro Asn Pro Leu Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly
     50                  55                  60

Pro Asp Pro Leu Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro
 65                  70                  75                  80

Ser Ala Asn Ile Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser
                 85                  90                  95

Asp Leu Tyr Gly Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly
            100                 105                 110

Pro Ser Gly Phe Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr
        115                 120                 125

Gly Glu Ser Ala Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu
    130                 135                 140
```

```
Ala Arg Tyr Val Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His
145                 150                 155                 160

Val Ser Trp His Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His
                165                 170                 175

Met Leu Leu Thr Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe
            180                 185                 190

Gly Val Val Thr Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu
        195                 200                 205

His Ser Ala Gln Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg
    210                 215                 220

Thr Val Pro Ile Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg
225                 230                 235                 240

Gly Glu Thr Ile Phe Glu Ile Asp Pro His Leu Gln Glu Arg Val Asp
                245                 250                 255

Lys Gly Ile Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala Lys
            260                 265                 270

Cys Ala Trp Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp Ser
        275                 280                 285

Arg Ser Ile Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys Asp
    290                 295                 300

Thr Glu Gln Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn Ser
305                 310                 315                 320

Lys Pro Val Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu Pro
                325                 330                 335

His Asp Arg Ala Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser Ser
            340                 345                 350

Thr Ala Ile Ile Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu Ser
        355                 360                 365

Val His Leu Lys Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu Cys
    370                 375                 380

Leu Arg Gly Arg Leu Leu His Gly Arg His Phe Thr Tyr Lys Ser Ile
385                 390                 395                 400
```

```
Thr Gly Asp Met Ala Ile Thr Phe Val Ser Thr Gly Val Glu Gly Ala
                405                 410                 415

Phe Ala Thr Glu Glu His Pro Tyr Ala Ala His Gly Pro Trp Leu Gln
            420                 425                 430

Ile Leu Leu Thr Glu Glu Phe Val Glu Lys Met Leu Glu Asp Leu Glu
        435                 440                 445

Asp Leu Thr Ser Pro Glu Glu Phe Lys Leu Pro Lys Glu Tyr Ser Trp
    450                 455                 460

Pro Glu Lys Lys Leu Lys Val Ser Ile Leu Pro Asp Val Val Phe Asp
465                 470                 475                 480

Ser Pro Leu His
```

<210> 2
<211> 105
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of human SUFU

<400> 2

```
Leu Ile Pro Leu Cys Leu Arg Gly Arg Leu Leu His Gly Arg His Phe
  1               5                  10                  15

Thr Tyr Lys Ser Ile Thr Gly Asp Met Ala Ile Thr Phe Val Ser Thr
             20                  25                  30

Gly Val Glu Gly Ala Phe Ala Thr Glu Glu His Pro Tyr Ala Ala His
         35                  40                  45

Gly Pro Trp Leu Gln Ile Leu Leu Thr Glu Glu Phe Val Glu Lys Met
     50                  55                  60

Leu Glu Asp Leu Glu Asp Leu Thr Ser Pro Glu Glu Phe Lys Leu Pro
 65                  70                  75                  80

Lys Glu Tyr Ser Trp Pro Glu Lys Lys Leu Lys Val Ser Ile Leu Pro
                 85                  90                  95

Asp Val Val Phe Asp Ser Pro Leu His
```

100 105

<210> 3
<211> 98
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of human SUFU

<400> 3
Gly Arg Leu Leu His Gly Arg His Phe Thr Tyr Lys Ser Ile Thr Gly
1 5 10 15

Asp Met Ala Ile Thr Phe Val Ser Thr Gly Val Glu Gly Ala Phe Ala
20 25 30

Thr Glu Glu His Pro Tyr Ala Ala His Gly Pro Trp Leu Gln Ile Leu
35 40 45

Leu Thr Glu Glu Phe Val Glu Lys Met Leu Glu Asp Leu Glu Asp Leu
50 55 60

Thr Ser Pro Glu Glu Phe Lys Leu Pro Lys Glu Tyr Ser Trp Pro Glu
65 70 75 80

Lys Lys Leu Lys Val Ser Ile Leu Pro Asp Val Val Phe Asp Ser Pro
85 90 95

Leu His

<210> 4
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of human SUFU

<400> 4
Gly Arg Leu Leu His Gly Arg His Phe Thr Tyr Lys Ser Ile Thr Gly
1 5 10 15

Asp Met Ala

<210> 5
<211> 1106
<212> PRT
<213> Homo sapiens

<400> 5

```
Met Phe Asn Ser Met Thr Pro Pro Pro Ile Ser Ser Tyr Gly Glu Pro
  1               5                  10                  15

Cys Cys Leu Arg Pro Leu Pro Ser Gln Gly Ala Pro Ser Val Gly Thr
             20                  25                  30

Glu Gly Leu Ser Gly Pro Pro Phe Cys His Gln Ala Asn Leu Met Ser
         35                  40                  45

Gly Pro His Ser Tyr Gly Pro Ala Arg Glu Thr Asn Ser Cys Thr Glu
     50                  55                  60

Gly Pro Leu Phe Ser Ser Pro Arg Ser Ala Val Lys Leu Thr Lys Lys
 65                  70                  75                  80

Arg Ala Leu Ser Ile Ser Pro Leu Ser Asp Ala Ser Leu Asp Leu Gln
                 85                  90                  95

Thr Val Ile Arg Thr Ser Pro Ser Ser Leu Val Ala Phe Ile Asn Ser
            100                 105                 110

Arg Cys Thr Ser Pro Gly Gly Ser Tyr Gly His Leu Ser Ile Gly Thr
        115                 120                 125

Met Ser Pro Ser Leu Gly Phe Pro Ala Gln Met Asn His Gln Lys Gly
    130                 135                 140

Pro Ser Pro Ser Phe Gly Val Gln Pro Cys Gly Pro His Asp Ser Ala
145                 150                 155                 160

Arg Gly Gly Met Ile Pro His Pro Gln Ser Arg Gly Pro Phe Pro Thr
                165                 170                 175

Cys Gln Leu Lys Ser Glu Leu Asp Met Leu Val Gly Lys Cys Arg Glu
            180                 185                 190

Glu Pro Leu Glu Gly Asp Met Ser Ser Pro Asn Ser Thr Gly Ile Gln
        195                 200                 205
```

```
Asp Pro Leu Leu Gly Met Leu Asp Gly Arg Glu Asp Leu Glu Arg Glu
    210                 215                 220

Glu Lys Arg Glu Pro Glu Ser Val Tyr Glu Thr Asp Cys Arg Trp Asp
225                 230                 235                 240

Gly Cys Ser Gln Glu Phe Asp Ser Gln Glu Gln Leu Val His His Ile
                245                 250                 255

Asn Ser Glu His Ile His Gly Glu Arg Lys Glu Phe Val Cys His Trp
            260                 265                 270

Gly Gly Cys Ser Arg Glu Leu Arg Pro Phe Lys Ala Gln Tyr Met Leu
        275                 280                 285

Val Val His Met Arg Arg His Thr Gly Glu Lys Pro His Lys Cys Thr
    290                 295                 300

Phe Glu Gly Cys Arg Lys Ser Tyr Ser Arg Leu Glu Asn Leu Lys Thr
305                 310                 315                 320

His Leu Arg Ser His Thr Gly Glu Lys Pro Tyr Met Cys Glu His Glu
                325                 330                 335

Gly Cys Ser Lys Ala Phe Ser Asn Ala Ser Asp Arg Ala Lys His Gln
            340                 345                 350

Asn Arg Thr His Ser Asn Glu Lys Pro Tyr Val Cys Lys Leu Pro Gly
        355                 360                 365

Cys Thr Lys Arg Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys
    370                 375                 380

Thr Val His Gly Pro Asp Ala His Val Thr Lys Arg His Arg Gly Asp
385                 390                 395                 400

Gly Pro Leu Pro Arg Ala Pro Ser Ile Ser Thr Val Glu Pro Lys Arg
                405                 410                 415

Glu Arg Glu Gly Gly Pro Ile Arg Glu Glu Ser Arg Leu Thr Val Pro
            420                 425                 430

Glu Gly Ala Met Lys Pro Gln Pro Ser Pro Gly Ala Gln Ser Ser Cys
        435                 440                 445

Ser Ser Asp His Ser Pro Ala Gly Ser Ala Ala Asn Thr Asp Ser Gly
    450                 455                 460
```

```
Val Glu Met Thr Gly Asn Ala Gly Gly Ser Thr Glu Asp Leu Ser Ser
465                 470                 475                 480

Leu Asp Glu Gly Pro Cys Ile Ala Gly Thr Gly Leu Ser Thr Leu Arg
                485                 490                 495

Arg Leu Glu Asn Leu Arg Leu Asp Gln Leu His Gln Leu Arg Pro Ile
            500                 505                 510

Gly Thr Arg Gly Leu Lys Leu Pro Ser Leu Ser His Thr Gly Thr Thr
        515                 520                 525

Val Ser Arg Arg Val Gly Pro Pro Val Ser Leu Glu Arg Arg Ser Ser
    530                 535                 540

Ser Ser Ser Ser Ile Ser Ser Ala Tyr Thr Val Ser Arg Arg Ser Ser
545                 550                 555                 560

Leu Ala Ser Pro Phe Pro Pro Gly Ser Pro Pro Glu Asn Gly Ala Ser
                565                 570                 575

Ser Leu Pro Gly Leu Met Pro Ala Gln His Tyr Leu Leu Arg Ala Arg
            580                 585                 590

Tyr Ala Ser Ala Arg Gly Gly Gly Thr Ser Pro Thr Ala Ala Ser Ser
        595                 600                 605

Leu Asp Arg Ile Gly Gly Leu Pro Met Pro Pro Trp Arg Ser Arg Ala
    610                 615                 620

Glu Tyr Pro Gly Tyr Asn Pro Asn Ala Gly Val Thr Arg Arg Ala Ser
625                 630                 635                 640

Asp Pro Ala Gln Ala Ala Asp Arg Pro Ala Pro Ala Arg Val Gln Arg
                645                 650                 655

Phe Lys Ser Leu Gly Cys Val His Thr Pro Pro Thr Val Ala Gly Gly
            660                 665                 670

Gly Gln Asn Phe Asp Pro Tyr Leu Pro Thr Ser Val Tyr Ser Pro Gln
        675                 680                 685

Pro Pro Ser Ile Thr Glu Asn Ala Ala Met Asp Ala Arg Gly Leu Gln
    690                 695                 700

Glu Glu Pro Glu Val Gly Thr Ser Met Val Gly Ser Gly Leu Asn Pro
705                 710                 715                 720
```

```
Tyr Met Asp Phe Pro Pro Thr Asp Thr Leu Gly Tyr Gly Gly Pro Glu
                725                 730                 735

Gly Ala Ala Ala Glu Pro Tyr Gly Ala Arg Gly Pro Gly Ser Leu Pro
            740                 745                 750

Leu Gly Pro Gly Pro Pro Thr Asn Tyr Gly Pro Asn Pro Cys Pro Gln
        755                 760                 765

Gln Ala Ser Tyr Pro Asp Pro Thr Gln Glu Thr Trp Gly Glu Phe Pro
    770                 775                 780

Ser His Ser Gly Leu Tyr Pro Gly Pro Lys Ala Leu Gly Gly Thr Tyr
785                 790                 795                 800

Ser Gln Cys Pro Arg Leu Glu His Tyr Gly Gln Val Gln Val Lys Pro
                805                 810                 815

Glu Gln Gly Cys Pro Val Gly Ser Asp Ser Thr Gly Leu Ala Pro Cys
            820                 825                 830

Leu Asn Ala His Pro Ser Glu Gly Pro Pro His Pro Gln Pro Leu Phe
        835                 840                 845

Ser His Tyr Pro Gln Pro Ser Pro Pro Gln Tyr Leu Gln Ser Gly Pro
    850                 855                 860

Tyr Thr Gln Pro Pro Pro Asp Tyr Leu Pro Ser Glu Pro Arg Pro Cys
865                 870                 875                 880

Leu Asp Phe Asp Ser Pro Thr His Ser Thr Gly Gln Leu Lys Ala Gln
                885                 890                 895

Leu Val Cys Asn Tyr Val Gln Ser Gln Gln Glu Leu Leu Trp Glu Gly
            900                 905                 910

Gly Gly Arg Glu Asp Ala Pro Ala Gln Glu Pro Ser Tyr Gln Ser Pro
        915                 920                 925

Lys Phe Leu Gly Gly Ser Gln Val Ser Pro Ser Arg Ala Lys Ala Pro
    930                 935                 940

Val Asn Thr Tyr Gly Pro Gly Phe Gly Pro Asn Leu Pro Asn His Lys
945                 950                 955                 960

Ser Gly Ser Tyr Pro Thr Pro Ser Pro Cys His Glu Asn Phe Val Val
                965                 970                 975
```

```
Gly Ala Asn Arg Ala Ser His Arg Ala Ala Ala Pro Pro Arg Leu Leu
            980                 985                 990

Pro Pro Leu Pro Thr Cys Tyr Gly Pro Leu Lys Val Gly Gly Thr Asn
        995                1000                1005

Pro Ser Cys Gly His Pro Glu Val Gly Arg Leu Gly Gly Gly Pro Ala
   1010                1015                1020

Leu Tyr Pro Pro Pro Glu Gly Gln Val Cys Asn Pro Leu Asp Ser Leu
1025                1030                1035                1040

Asp Leu Asp Asn Thr Gln Leu Asp Phe Val Ala Ile Leu Asp Glu Pro
               1045                1050                1055

Gln Gly Leu Ser Pro Pro Pro Ser His Asp Gln Arg Gly Ser Ser Gly
           1060                1065                1070

His Thr Pro Pro Pro Ser Gly Pro Pro Asn Met Ala Val Gly Asn Met
       1075                1080                1085

Ser Val Leu Leu Arg Ser Leu Pro Gly Glu Thr Glu Phe Leu Asn Ser
   1090                1095                1100

Ser Ala
1105
```

<210> 6
<211> 407
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of
human GLI-1

<400> 6

```
Met Phe Asn Ser Met Thr Pro Pro Pro Ile Ser Ser Tyr Gly Glu Pro
  1               5                  10                  15

Cys Cys Leu Arg Pro Leu Pro Ser Gln Gly Ala Pro Ser Val Gly Thr
             20                  25                  30

Glu Gly Leu Ser Gly Pro Pro Phe Cys His Gln Ala Asn Leu Met Ser
         35                  40                  45
```

```
Gly Pro His Ser Tyr Gly Pro Ala Arg Glu Thr Asn Ser Cys Thr Glu
     50                  55                  60

Gly Pro Leu Phe Ser Ser Pro Arg Ser Ala Val Lys Leu Thr Lys Lys
 65                  70                  75                  80

Arg Ala Leu Ser Ile Ser Pro Leu Ser Asp Ala Ser Leu Asp Leu Gln
                 85                  90                  95

Thr Val Ile Arg Thr Ser Pro Ser Ser Leu Val Ala Phe Ile Asn Ser
            100                 105                 110

Arg Cys Thr Ser Pro Gly Gly Ser Tyr Gly His Leu Ser Ile Gly Thr
        115                 120                 125

Met Ser Pro Ser Leu Gly Phe Pro Ala Gln Met Asn His Gln Lys Gly
    130                 135                 140

Pro Ser Pro Ser Phe Gly Val Gln Pro Cys Gly Pro His Asp Ser Ala
145                 150                 155                 160

Arg Gly Gly Met Ile Pro His Pro Gln Ser Arg Gly Pro Phe Pro Thr
                165                 170                 175

Cys Gln Leu Lys Ser Glu Leu Asp Met Leu Val Gly Lys Cys Arg Glu
            180                 185                 190

Glu Pro Leu Glu Gly Asp Met Ser Ser Pro Asn Ser Thr Gly Ile Gln
        195                 200                 205

Asp Pro Leu Leu Gly Met Leu Asp Gly Arg Glu Asp Leu Glu Arg Glu
    210                 215                 220

Glu Lys Arg Glu Pro Glu Ser Val Tyr Glu Thr Asp Cys Arg Trp Asp
225                 230                 235                 240

Gly Cys Ser Gln Glu Phe Asp Ser Gln Glu Gln Leu Val His His Ile
                245                 250                 255

Asn Ser Glu His Ile His Gly Glu Arg Lys Glu Phe Val Cys His Trp
            260                 265                 270

Gly Gly Cys Ser Arg Glu Leu Arg Pro Phe Lys Ala Gln Tyr Met Leu
        275                 280                 285

Val Val His Met Arg Arg His Thr Gly Glu Lys Pro His Lys Cys Thr
    290                 295                 300
```

```
Phe Glu Gly Cys Arg Lys Ser Tyr Ser Arg Leu Glu Asn Leu Lys Thr
305                 310                 315                 320

His Leu Arg Ser His Thr Gly Glu Lys Pro Tyr Met Cys Glu His Glu
                325                 330                 335

Gly Cys Ser Lys Ala Phe Ser Asn Ala Ser Asp Arg Ala Lys His Gln
            340                 345                 350

Asn Arg Thr His Ser Asn Glu Lys Pro Tyr Val Cys Lys Leu Pro Gly
        355                 360                 365

Cys Thr Lys Arg Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys
    370                 375                 380

Thr Val His Gly Pro Asp Ala His Val Thr Lys Arg His Arg Gly Asp
385                 390                 395                 400

Gly Pro Leu Pro Arg Ala Pro
                405
```

<210> 7
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of human GLI-1

<400> 7

```
Lys His Val Lys Thr Val His Gly Pro Asp Ala His Val Thr Lys Arg
  1               5                  10                  15

His Arg Gly Asp Gly Pro Leu Pro Arg Ala Pro
             20                  25
```

<210> 8
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of human GLI-1

<400> 8
Cys Thr Glu Gly Pro Leu Phe Ser Ser Pro Arg Ser Ala Val Lys Leu
1 5 10 15

Thr Lys Lys Arg Ala
20

<210> 9
<211> 468
<212> PRT
<213> Drosophila melanogaster

<400> 9
Met Ala Glu Ala Asn Leu Asp Lys Lys Pro Glu Val Lys Pro Pro Pro
1 5 10 15

Gly Leu Lys Ala Ile Ile Asp His Leu Gly Gln Val Tyr Pro Asn Gln
20 25 30

Pro Asn Pro Leu Gln Val Thr Thr Leu Leu Lys Tyr Trp Leu Gly Gly
35 40 45

Gln Asp Pro Leu Asp Tyr Ile Ser Met Tyr Lys Phe Pro Gly Asp Val
50 55 60

Asp Arg Asn Val Pro Pro His Trp His Tyr Ile Ser Phe Gly Leu Ser
65 70 75 80

Asp Leu His Gly Asp Glu Arg Val His Leu Arg Glu Glu Gly Val Thr
85 90 95

Arg Ser Gly Met Gly Phe Glu Leu Thr Phe Arg Leu Ala Lys Thr Glu
100 105 110

Ile Glu Leu Lys Gln Gln Ile Glu Asn Pro Glu Lys Pro Gln Arg Ala
115 120 125

Pro Thr Trp Pro Ala Asn Leu Leu Gln Ala Ile Gly Arg Tyr Cys Phe
130 135 140

Gln Thr Gly Asn Gly Leu Cys Phe Gly Asp Asn Ile Pro Trp Arg Lys
145 150 155 160

Ser Leu Asp Gly Ser Thr Thr Ser Lys Leu Gln Asn Leu Leu Val Ala
165 170 175

Gln Asp Pro Gln Leu Gly Cys Ile Asp Thr Pro Thr Gly Thr Val Asp

```
            180                 185                 190

Phe Cys Gln Ile Val Gly Val Phe Asp Asp Glu Leu Glu Gln Ala Ser
        195                 200                 205

Arg Trp Asn Gly Arg Gly Val Leu Asn Phe Leu Arg Gln Asp Met Gln
    210                 215                 220

Thr Gly Gly Asp Trp Leu Val Thr Asn Met Asp Arg Gln Met Ser Val
225                 230                 235                 240

Phe Glu Leu Phe Pro Glu Thr Leu Leu Asn Leu Gln Asp Asp Leu Glu
                245                 250                 255

Lys Gln Gly Ser Asp Leu Ala Gly Val Asn Ala Asp Phe Ser Phe Arg
            260                 265                 270

Glu Leu Lys Pro Thr Lys Glu Val Lys Glu Glu Val Asp Phe Gln Ala
        275                 280                 285

Leu Ser Glu Lys Cys Ala Asn Asp Glu Asn Asn Arg Gln Leu Thr Asp
    290                 295                 300

Thr Gln Met Lys Arg Glu Glu Pro Ser Phe Pro Gln Ser Met Ser Met
305                 310                 315                 320

Ser Ser Asn Ser Leu His Lys Ser Cys Pro Leu Asp Phe Gln Ala Gln
                325                 330                 335

Ala Pro Asn Cys Ile Ser Leu Asp Gly Ile Glu Ile Thr Leu Ala Pro
            340                 345                 350

Gly Val Ala Lys Tyr Leu Leu Leu Ala Ile Lys Asp Arg Ile Arg His
        355                 360                 365

Gly Arg His Phe Thr Phe Lys Ala Gln His Leu Ala Leu Thr Leu Val
    370                 375                 380

Ala Glu Ser Val Thr Gly Ser Ala Val Thr Val Asn Glu Pro Tyr Gly
385                 390                 395                 400

Val Leu Gly Tyr Trp Ile Gln Val Leu Ile Pro Asp Glu Leu Val Pro
                405                 410                 415

Arg Leu Met Glu Asp Phe Cys Ser Ala Gly Leu Asp Glu Lys Cys Glu
            420                 425                 430

Pro Lys Glu Arg Leu Glu Leu Glu Trp Pro Asp Lys Asn Leu Lys Leu
```

```
        435                 440                 445

Ile Ile Asp Gln Pro Glu Pro Val Leu Pro Met Ser Leu Asp Ala Ala
    450                 455                 460

Pro Leu Lys Met
465
```

<210> 10
<211> 384
<212> PRT
<213> Homo sapiens

<400> 10

```
Met Ala Glu Leu Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro
  1               5                  10                  15

Ala Pro Gly Pro Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro
             20                  25                  30

Gly Leu His Ala Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln
         35                  40                  45

Pro Asn Pro Leu Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly
     50                  55                  60

Pro Asp Pro Leu Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro
 65                  70                  75                  80

Ser Ala Asn Ile Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser
                 85                  90                  95

Asp Leu Tyr Gly Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly
            100                 105                 110

Pro Ser Gly Phe Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr
        115                 120                 125

Gly Glu Ser Ala Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu
    130                 135                 140

Ala Arg Tyr Val Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His
145                 150                 155                 160

Val Ser Trp His Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His
                165                 170                 175
```

```
Met Leu Leu Thr Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe
            180                 185                 190

Gly Val Val Thr Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu
        195                 200                 205

His Ser Ala Gln Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg
    210                 215                 220

Thr Val Pro Ile Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg
225                 230                 235                 240

Gly Glu Thr Ile Phe Glu Ile Asp Pro His Leu Gln Glu Arg Val Asp
                245                 250                 255

Lys Gly Ile Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala Lys
            260                 265                 270

Cys Ala Trp Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp Ser
        275                 280                 285

Arg Ser Ile Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys Asp
    290                 295                 300

Thr Glu Gln Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn Ser
305                 310                 315                 320

Lys Pro Val Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu Pro
                325                 330                 335

His Asp Arg Ala Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser Ser
            340                 345                 350

Thr Ala Ile Ile Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu Ser
        355                 360                 365

Val His Leu Lys Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu Cys
    370                 375                 380
```

<210> 11
<211> 1118
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Haemoagglutinin-tagged GLI-1

<400> 11

```
Met Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Met Phe Asn Ser
  1               5                  10                  15

Met Thr Pro Pro Pro Ile Ser Ser Tyr Gly Glu Pro Cys Cys Leu Arg
             20                  25                  30

Pro Leu Pro Ser Gln Gly Ala Pro Ser Val Gly Thr Glu Gly Leu Ser
         35                  40                  45

Gly Pro Pro Phe Cys His Gln Ala Asn Leu Met Ser Gly Pro His Ser
     50                  55                  60

Tyr Gly Pro Ala Arg Glu Thr Asn Ser Cys Thr Glu Gly Pro Leu Phe
 65                  70                  75                  80

Ser Ser Pro Arg Ser Ala Val Lys Leu Thr Lys Lys Arg Ala Leu Ser
                 85                  90                  95

Ile Ser Pro Leu Ser Asp Ala Ser Leu Asp Leu Gln Thr Val Ile Arg
            100                 105                 110

Thr Ser Pro Ser Ser Leu Val Ala Phe Ile Asn Ser Arg Cys Thr Ser
        115                 120                 125

Pro Gly Gly Ser Tyr Gly His Leu Ser Ile Gly Thr Met Ser Pro Ser
    130                 135                 140

Leu Gly Phe Pro Ala Gln Met Asn His Gln Lys Gly Pro Ser Pro Ser
145                 150                 155                 160

Phe Gly Val Gln Pro Cys Gly Pro His Asp Ser Ala Arg Gly Gly Met
                165                 170                 175

Ile Pro His Pro Gln Ser Arg Gly Pro Phe Pro Thr Cys Gln Leu Lys
            180                 185                 190

Ser Glu Leu Asp Met Leu Val Gly Lys Cys Arg Glu Glu Pro Leu Glu
        195                 200                 205

Gly Asp Met Ser Ser Pro Asn Ser Thr Gly Ile Gln Asp Pro Leu Leu
    210                 215                 220

Gly Met Leu Asp Gly Arg Glu Asp Leu Glu Arg Glu Glu Lys Arg Glu
```

225 230 235 240

Pro Glu Ser Val Tyr Glu Thr Asp Cys Arg Trp Asp Gly Cys Ser Gln
245 250 255

Glu Phe Asp Ser Gln Glu Gln Leu Val His His Ile Asn Ser Glu His
260 265 270

Ile His Gly Glu Arg Lys Glu Phe Val Cys His Trp Gly Gly Cys Ser
275 280 285

Arg Glu Leu Arg Pro Phe Lys Ala Gln Tyr Met Leu Val Val His Met
290 295 300

Arg Arg His Thr Gly Glu Lys Pro His Lys Cys Thr Phe Glu Gly Cys
305 310 315 320

Arg Lys Ser Tyr Ser Arg Leu Glu Asn Leu Lys Thr His Leu Arg Ser
325 330 335

His Thr Gly Glu Lys Pro Tyr Met Cys Glu His Glu Gly Cys Ser Lys
340 345 350

Ala Phe Ser Asn Ala Ser Asp Arg Ala Lys His Gln Asn Arg Thr His
355 360 365

Ser Asn Glu Lys Pro Tyr Val Cys Lys Leu Pro Gly Cys Thr Lys Arg
370 375 380

Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys Thr Val His Gly
385 390 395 400

Pro Asp Ala His Val Thr Lys Arg His Arg Gly Asp Gly Pro Leu Pro
405 410 415

Arg Ala Pro Ser Ile Ser Thr Val Glu Pro Lys Arg Glu Arg Glu Gly
420 425 430

Gly Pro Ile Arg Glu Glu Ser Arg Leu Thr Val Pro Glu Gly Ala Met
435 440 445

Lys Pro Gln Pro Ser Pro Gly Ala Gln Ser Ser Cys Ser Ser Asp His
450 455 460

Ser Pro Ala Gly Ser Ala Ala Asn Thr Asp Ser Gly Val Glu Met Thr
465 470 475 480

Gly Asn Ala Gly Gly Ser Thr Glu Asp Leu Ser Ser Leu Asp Glu Gly

```
                485                 490                 495

Pro Cys Ile Ala Gly Thr Gly Leu Ser Thr Leu Arg Arg Leu Glu Asn
            500                 505                 510

Leu Arg Leu Asp Gln Leu His Gln Leu Arg Pro Ile Gly Thr Arg Gly
        515                 520                 525

Leu Lys Leu Pro Ser Leu Ser His Thr Gly Thr Thr Val Ser Arg Arg
    530                 535                 540

Val Gly Pro Pro Val Ser Leu Glu Arg Arg Ser Ser Ser Ser Ser Ser
545                 550                 555                 560

Ile Ser Ser Ala Tyr Thr Val Ser Arg Arg Ser Ser Leu Ala Ser Pro
                565                 570                 575

Phe Pro Pro Gly Ser Pro Pro Glu Asn Gly Ala Ser Ser Leu Pro Gly
            580                 585                 590

Leu Met Pro Ala Gln His Tyr Leu Leu Arg Ala Arg Tyr Ala Ser Ala
        595                 600                 605

Arg Gly Gly Gly Thr Ser Pro Thr Ala Ala Ser Ser Leu Asp Arg Ile
    610                 615                 620

Gly Gly Leu Pro Met Pro Pro Trp Arg Ser Arg Ala Glu Tyr Pro Gly
625                 630                 635                 640

Tyr Asn Pro Asn Ala Gly Val Thr Arg Arg Ala Ser Asp Pro Ala Gln
                645                 650                 655

Ala Ala Asp Arg Pro Ala Pro Ala Arg Val Gln Arg Phe Lys Ser Leu
            660                 665                 670

Gly Cys Val His Thr Pro Pro Thr Val Ala Gly Gly Gly Gln Asn Phe
        675                 680                 685

Asp Pro Tyr Leu Pro Thr Ser Val Tyr Ser Pro Gln Pro Pro Ser Ile
    690                 695                 700

Thr Glu Asn Ala Ala Met Asp Ala Arg Gly Leu Gln Glu Glu Pro Glu
705                 710                 715                 720

Val Gly Thr Ser Met Val Gly Ser Gly Leu Asn Pro Tyr Met Asp Phe
                725                 730                 735

Pro Pro Thr Asp Thr Leu Gly Tyr Gly Gly Pro Glu Gly Ala Ala Ala
```

```
            740                 745                 750

Glu Pro Tyr Gly Ala Arg Gly Pro Gly Ser Leu Pro Leu Gly Pro Gly
        755                 760                 765

Pro Pro Thr Asn Tyr Gly Pro Asn Pro Cys Pro Gln Gln Ala Ser Tyr
    770                 775                 780

Pro Asp Pro Thr Gln Glu Thr Trp Gly Glu Phe Pro Ser His Ser Gly
785                 790                 795                 800

Leu Tyr Pro Gly Pro Lys Ala Leu Gly Gly Thr Tyr Ser Gln Cys Pro
                805                 810                 815

Arg Leu Glu His Tyr Gly Gln Val Gln Val Lys Pro Glu Gln Gly Cys
            820                 825                 830

Pro Val Gly Ser Asp Ser Thr Gly Leu Ala Pro Cys Leu Asn Ala His
        835                 840                 845

Pro Ser Glu Gly Pro Pro His Pro Gln Pro Leu Phe Ser His Tyr Pro
    850                 855                 860

Gln Pro Ser Pro Pro Gln Tyr Leu Gln Ser Gly Pro Tyr Thr Gln Pro
865                 870                 875                 880

Pro Pro Asp Tyr Leu Pro Ser Glu Pro Arg Pro Cys Leu Asp Phe Asp
                885                 890                 895

Ser Pro Thr His Ser Thr Gly Gln Leu Lys Ala Gln Leu Val Cys Asn
            900                 905                 910

Tyr Val Gln Ser Gln Gln Glu Leu Leu Trp Glu Gly Gly Gly Arg Glu
        915                 920                 925

Asp Ala Pro Ala Gln Glu Pro Ser Tyr Gln Ser Pro Lys Phe Leu Gly
    930                 935                 940

Gly Ser Gln Val Ser Pro Ser Arg Ala Lys Ala Pro Val Asn Thr Tyr
945                 950                 955                 960

Gly Pro Gly Phe Gly Pro Asn Leu Pro Asn His Lys Ser Gly Ser Tyr
                965                 970                 975

Pro Thr Pro Ser Pro Cys His Glu Asn Phe Val Val Gly Ala Asn Arg
            980                 985                 990

Ala Ser His Arg Ala Ala Ala Pro Pro Arg Leu Leu Pro Pro Leu Pro
```

995 1000 1005

Thr Cys Tyr Gly Pro Leu Lys Val Gly Gly Thr Asn Pro Ser Cys Gly
1010 1015 1020

His Pro Glu Val Gly Arg Leu Gly Gly Gly Pro Ala Leu Tyr Pro Pro
1025 1030 1035 1040

Pro Glu Gly Gln Val Cys Asn Pro Leu Asp Ser Leu Asp Leu Asp Asn
1045 1050 1055

Thr Gln Leu Asp Phe Val Ala Ile Leu Asp Glu Pro Gln Gly Leu Ser
1060 1065 1070

Pro Pro Pro Ser His Asp Gln Arg Gly Ser Ser Gly His Thr Pro Pro
1075 1080 1085

Pro Ser Gly Pro Pro Asn Met Ala Val Gly Asn Met Ser Val Leu Leu
1090 1095 1100

Arg Ser Leu Pro Gly Glu Thr Glu Phe Leu Asn Ser Ser Ala
1105 1110 1115

<210> 12
<211> 496
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:myc-tagged
SUFUH

<400> 12
Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met Ala Glu Leu
1 5 10 15

Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro Ala Pro Gly Pro
20 25 30

Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro Gly Leu His Ala
35 40 45

Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln Pro Asn Pro Leu
50 55 60

Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly Pro Asp Pro Leu
65 70 75 80

```
Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro Ser Ala Asn Ile
                 85                  90                  95

Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser Asp Leu Tyr Gly
            100                 105                 110

Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly Pro Ser Gly Phe
        115                 120                 125

Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr Gly Glu Ser Ala
    130                 135                 140

Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu Ala Arg Tyr Val
145                 150                 155                 160

Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His Val Ser Trp His
                165                 170                 175

Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His Met Leu Leu Thr
            180                 185                 190

Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe Gly Val Val Thr
        195                 200                 205

Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu His Ser Ala Gln
    210                 215                 220

Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg Thr Val Pro Ile
225                 230                 235                 240

Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg Gly Glu Thr Ile
                245                 250                 255

Phe Glu Ile Asp Pro His Leu Gln Glu Arg Val Asp Lys Gly Ile Glu
            260                 265                 270

Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala Lys Cys Ala Trp Asp
        275                 280                 285

Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp Ser Arg Ser Ile Cys
    290                 295                 300

Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys Asp Thr Glu Gln Ile
305                 310                 315                 320

Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn Ser Lys Pro Val Leu
                325                 330                 335
```

```
Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu Pro His Asp Arg Ala
            340                 345                 350

Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser Ser Thr Ala Ile Ile
        355                 360                 365

Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu Ser Val His Leu Lys
    370                 375                 380

Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu Cys Leu Arg Gly Arg
385                 390                 395                 400

Leu Leu His Gly Arg His Phe Thr Tyr Lys Ser Ile Thr Gly Asp Met
                405                 410                 415

Ala Ile Thr Phe Val Ser Thr Gly Val Glu Gly Ala Phe Ala Thr Glu
            420                 425                 430

Glu His Pro Tyr Ala Ala His Gly Pro Trp Leu Gln Ile Leu Leu Thr
        435                 440                 445

Glu Glu Phe Val Glu Lys Met Leu Glu Asp Leu Glu Asp Leu Thr Ser
    450                 455                 460

Pro Glu Glu Phe Lys Leu Pro Lys Glu Tyr Ser Trp Pro Glu Lys Lys
465                 470                 475                 480

Leu Lys Val Ser Ile Leu Pro Asp Val Val Phe Asp Ser Pro Leu His
                485                 490                 495
```

<210> 13
<211> 419
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial
Sequence:Haemoagglutinin-tagged fragment of human
GLI-1

```
<400> 13
Met Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Met Phe Asn Ser
  1               5                  10                  15
```

```
Met Thr Pro Pro Pro Ile Ser Ser Tyr Gly Glu Pro Cys Cys Leu Arg
            20                  25                  30

Pro Leu Pro Ser Gln Gly Ala Pro Ser Val Gly Thr Glu Gly Leu Ser
        35                  40                  45

Gly Pro Pro Phe Cys His Gln Ala Asn Leu Met Ser Gly Pro His Ser
    50                  55                  60

Tyr Gly Pro Ala Arg Glu Thr Asn Ser Cys Thr Glu Gly Pro Leu Phe
65                  70                  75                  80

Ser Ser Pro Arg Ser Ala Val Lys Leu Thr Lys Lys Arg Ala Leu Ser
                85                  90                  95

Ile Ser Pro Leu Ser Asp Ala Ser Leu Asp Leu Gln Thr Val Ile Arg
            100                 105                 110

Thr Ser Pro Ser Ser Leu Val Ala Phe Ile Asn Ser Arg Cys Thr Ser
        115                 120                 125

Pro Gly Gly Ser Tyr Gly His Leu Ser Ile Gly Thr Met Ser Pro Ser
    130                 135                 140

Leu Gly Phe Pro Ala Gln Met Asn His Gln Lys Gly Pro Ser Pro Ser
145                 150                 155                 160

Phe Gly Val Gln Pro Cys Gly Pro His Asp Ser Ala Arg Gly Gly Met
                165                 170                 175

Ile Pro His Pro Gln Ser Arg Gly Pro Phe Pro Thr Cys Gln Leu Lys
            180                 185                 190

Ser Glu Leu Asp Met Leu Val Gly Lys Cys Arg Glu Glu Pro Leu Glu
        195                 200                 205

Gly Asp Met Ser Ser Pro Asn Ser Thr Gly Ile Gln Asp Pro Leu Leu
    210                 215                 220

Gly Met Leu Asp Gly Arg Glu Asp Leu Glu Arg Glu Glu Lys Arg Glu
225                 230                 235                 240

Pro Glu Ser Val Tyr Glu Thr Asp Cys Arg Trp Asp Gly Cys Ser Gln
                245                 250                 255

Glu Phe Asp Ser Gln Glu Gln Leu Val His His Ile Asn Ser Glu His
            260                 265                 270
```

```
Ile His Gly Glu Arg Lys Glu Phe Val Cys His Trp Gly Gly Cys Ser
        275                 280                 285

Arg Glu Leu Arg Pro Phe Lys Ala Gln Tyr Met Leu Val Val His Met
    290                 295                 300

Arg Arg His Thr Gly Glu Lys Pro His Lys Cys Thr Phe Glu Gly Cys
305                 310                 315                 320

Arg Lys Ser Tyr Ser Arg Leu Glu Asn Leu Lys Thr His Leu Arg Ser
                325                 330                 335

His Thr Gly Glu Lys Pro Tyr Met Cys Glu His Glu Gly Cys Ser Lys
            340                 345                 350

Ala Phe Ser Asn Ala Ser Asp Arg Ala Lys His Gln Asn Arg Thr His
        355                 360                 365

Ser Asn Glu Lys Pro Tyr Val Cys Lys Leu Pro Gly Cys Thr Lys Arg
    370                 375                 380

Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys Thr Val His Gly
385                 390                 395                 400

Pro Asp Ala His Val Thr Lys Arg His Arg Gly Asp Gly Pro Leu Pro
                405                 410                 415

Arg Ala Pro
```

<210> 14
<211> 397
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:myc-tagged SUFUH-delta

<400> 14

```
Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met Ala Glu Leu
  1               5                  10                  15

Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro Ala Pro Gly Pro
            20                  25                  30
```

```
Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro Gly Leu His Ala
         35                  40                  45

Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln Pro Asn Pro Leu
     50                  55                  60

Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly Pro Asp Pro Leu
 65                  70                  75                  80

Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro Ser Ala Asn Ile
                 85                  90                  95

Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser Asp Leu Tyr Gly
            100                 105                 110

Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly Pro Ser Gly Phe
        115                 120                 125

Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr Gly Glu Ser Ala
    130                 135                 140

Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu Ala Arg Tyr Val
145                 150                 155                 160

Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His Val Ser Trp His
                165                 170                 175

Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His Met Leu Leu Thr
            180                 185                 190

Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe Gly Val Val Thr
        195                 200                 205

Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu His Ser Ala Gln
    210                 215                 220

Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg Thr Val Pro Ile
225                 230                 235                 240

Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg Gly Glu Thr Ile
                245                 250                 255

Phe Glu Ile Asp Pro His Leu Gln Glu Arg Val Asp Lys Gly Ile Glu
            260                 265                 270

Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala Lys Cys Ala Trp Asp
        275                 280                 285
```

```
Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp Ser Arg Ser Ile Cys
    290                 295                 300

Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys Asp Thr Glu Gln Ile
305                 310                 315                 320

Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn Ser Lys Pro Val Leu
                325                 330                 335

Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu Pro His Asp Arg Ala
            340                 345                 350

Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser Ser Thr Ala Ile Ile
        355                 360                 365

Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu Ser Val His Leu Lys
    370                 375                 380

Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu Cys Leu
385                 390                 395
```

<210> 15
<211> 525
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of human GLI-1

<400> 15

```
Met Phe Asn Ser Met Thr Pro Pro Pro Ile Ser Ser Tyr Gly Glu Pro
  1               5                  10                  15

Cys Cys Leu Arg Pro Leu Pro Ser Gln Gly Ala Pro Ser Val Gly Thr
             20                  25                  30

Glu Gly Leu Ser Gly Pro Pro Phe Cys His Gln Ala Asn Leu Met Ser
         35                  40                  45

Gly Pro His Ser Tyr Gly Pro Ala Arg Glu Thr Asn Ser Cys Thr Glu
     50                  55                  60

Gly Pro Leu Phe Ser Ser Pro Arg Ser Ala Val Lys Leu Thr Lys Lys
 65                  70                  75                  80

Arg Ala Leu Ser Ile Ser Pro Leu Ser Asp Ala Ser Leu Asp Leu Gln
```

```
                85                  90                  95

Thr Val Ile Arg Thr Ser Pro Ser Ser Leu Val Ala Phe Ile Asn Ser
            100                 105                 110

Arg Cys Thr Ser Pro Gly Gly Ser Tyr Gly His Leu Ser Ile Gly Thr
        115                 120                 125

Met Ser Pro Ser Leu Gly Phe Pro Ala Gln Met Asn His Gln Lys Gly
    130                 135                 140

Pro Ser Pro Ser Phe Gly Val Gln Pro Cys Gly Pro His Asp Ser Ala
145                 150                 155                 160

Arg Gly Gly Met Ile Pro His Pro Gln Ser Arg Gly Pro Phe Pro Thr
                165                 170                 175

Cys Gln Leu Lys Ser Glu Leu Asp Met Leu Val Gly Lys Cys Arg Glu
            180                 185                 190

Glu Pro Leu Glu Gly Asp Met Ser Ser Pro Asn Ser Thr Gly Ile Gln
        195                 200                 205

Asp Pro Leu Leu Gly Met Leu Asp Gly Arg Glu Asp Leu Glu Arg Glu
    210                 215                 220

Glu Lys Arg Glu Pro Glu Ser Val Tyr Glu Thr Asp Cys Arg Trp Asp
225                 230                 235                 240

Gly Cys Ser Gln Glu Phe Asp Ser Gln Glu Gln Leu Val His His Ile
                245                 250                 255

Asn Ser Glu His Ile His Gly Glu Arg Lys Glu Phe Val Cys His Trp
            260                 265                 270

Gly Gly Cys Ser Arg Glu Leu Arg Pro Phe Lys Ala Gln Tyr Met Leu
        275                 280                 285

Val Val His Met Arg Arg His Thr Gly Glu Lys Pro His Lys Cys Thr
    290                 295                 300

Phe Glu Gly Cys Arg Lys Ser Tyr Ser Arg Leu Glu Asn Leu Lys Thr
305                 310                 315                 320

His Leu Arg Ser His Thr Gly Glu Lys Pro Tyr Met Cys Glu His Glu
                325                 330                 335

Gly Cys Ser Lys Ala Phe Ser Asn Ala Ser Asp Arg Ala Lys His Gln
```

```
            340                 345                 350

Asn Arg Thr His Ser Asn Glu Lys Pro Tyr Val Cys Lys Leu Pro Gly
        355                 360                 365

Cys Thr Lys Arg Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys
    370                 375                 380

Thr Val His Gly Pro Asp Ala His Val Thr Lys Arg His Arg Gly Asp
385                 390                 395                 400

Gly Pro Leu Pro Arg Ala Pro Ser Ile Ser Thr Val Glu Pro Lys Arg
                405                 410                 415

Glu Arg Glu Gly Gly Pro Ile Arg Glu Glu Ser Arg Leu Thr Val Pro
            420                 425                 430

Glu Gly Ala Met Lys Pro Gln Pro Ser Pro Gly Ala Gln Ser Ser Cys
        435                 440                 445

Ser Ser Asp His Ser Pro Ala Gly Ser Ala Ala Asn Thr Asp Ser Gly
    450                 455                 460

Val Glu Met Thr Gly Asn Ala Gly Gly Ser Thr Glu Asp Leu Ser Ser
465                 470                 475                 480

Leu Asp Glu Gly Pro Cys Ile Ala Gly Thr Gly Leu Ser Thr Leu Arg
                485                 490                 495

Arg Leu Glu Asn Leu Arg Leu Asp Gln Leu His Gln Leu Arg Pro Ile
            500                 505                 510

Gly Thr Arg Gly Leu Lys Leu Pro Ser Leu Ser His Thr
        515                 520                 525
```

<210> 16
<211> 525
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Fragment of the wild type of human GLI-1

<400> 16

```
Met Phe Asn Ser Met Thr Pro Pro Pro Ile Ser Ser Tyr Gly Glu Pro
  1               5                  10                  15
```

```
Cys Cys Leu Arg Pro Leu Pro Ser Gln Gly Ala Pro Ser Val Gly Thr
            20                  25                  30

Glu Gly Leu Ser Gly Pro Pro Phe Cys His Gln Ala Asn Leu Met Ser
        35                  40                  45

Gly Pro His Ser Tyr Gly Pro Ala Arg Glu Thr Asn Ser Cys Thr Glu
    50                  55                  60

Gly Pro Leu Phe Ser Ser Pro Arg Ser Ala Val Lys Leu Thr Lys Lys
65                  70                  75                  80

Arg Ala Leu Ser Ile Ser Pro Leu Ser Asp Ala Ser Leu Asp Leu Gln
                85                  90                  95

Thr Val Ile Arg Thr Ser Pro Ser Ser Leu Val Ala Phe Ile Asn Ser
            100                 105                 110

Arg Cys Thr Ser Pro Gly Gly Ser Tyr Gly His Leu Ser Ile Gly Thr
        115                 120                 125

Met Ser Pro Ser Leu Gly Phe Pro Ala Gln Met Asn His Gln Lys Gly
    130                 135                 140

Pro Ser Pro Ser Phe Gly Val Gln Pro Cys Gly Pro His Asp Ser Ala
145                 150                 155                 160

Arg Gly Gly Met Ile Pro His Pro Gln Ser Arg Gly Pro Phe Pro Thr
                165                 170                 175

Cys Gln Leu Lys Ser Glu Leu Asp Met Leu Val Gly Lys Cys Arg Glu
            180                 185                 190

Glu Pro Leu Glu Gly Asp Met Ser Ser Pro Asn Ser Thr Gly Ile Gln
        195                 200                 205

Asp Pro Leu Leu Gly Met Leu Asp Gly Arg Glu Asp Leu Glu Arg Glu
    210                 215                 220

Glu Lys Arg Glu Pro Glu Ser Val Tyr Glu Thr Asp Cys Arg Trp Asp
225                 230                 235                 240

Gly Cys Ser Gln Glu Phe Asp Ser Gln Glu Gln Leu Val His His Ile
                245                 250                 255

Asn Ser Glu His Ile His Gly Glu Arg Lys Glu Phe Val Cys His Trp
            260                 265                 270
```

```
Gly Gly Cys Ser Arg Glu Leu Arg Pro Phe Lys Ala Gln Tyr Met Leu
        275                 280                 285

Val Val His Met Arg Arg His Thr Gly Glu Lys Pro His Lys Cys Thr
    290                 295                 300

Phe Glu Gly Cys Arg Lys Ser Tyr Ser Arg Leu Glu Asn Leu Lys Thr
305                 310                 315                 320

His Leu Arg Ser His Thr Gly Glu Lys Pro Tyr Met Cys Glu His Glu
                325                 330                 335

Gly Cys Ser Lys Ala Phe Ser Asn Ala Ser Asp Arg Ala Lys His Gln
            340                 345                 350

Asn Arg Thr His Ser Asn Glu Lys Pro Tyr Val Cys Lys Leu Pro Gly
        355                 360                 365

Cys Thr Lys Arg Tyr Thr Asp Pro Ser Ser Leu Arg Lys His Val Lys
    370                 375                 380

Thr Val His Gly Pro Asp Ala His Val Thr Lys Arg His Arg Gly Asp
385                 390                 395                 400

Gly Pro Leu Pro Arg Ala Pro Ser Ile Ser Thr Val Glu Pro Lys Arg
                405                 410                 415

Glu Arg Glu Gly Gly Pro Ile Arg Glu Glu Ser Arg Leu Thr Val Pro
            420                 425                 430

Glu Gly Ala Met Lys Pro Gln Pro Ser Pro Gly Ala Gln Ser Ser Cys
        435                 440                 445

Ser Ser Asp His Ser Pro Ala Gly Ser Ala Ala Asn Thr Asp Ser Gly
    450                 455                 460

Val Glu Met Thr Gly Asn Ala Gly Gly Ser Thr Glu Asp Leu Ser Ser
465                 470                 475                 480

Leu Asp Glu Gly Pro Cys Ile Ala Gly Thr Gly Leu Ser Thr Ala Arg
                485                 490                 495

Arg Leu Glu Asn Ala Arg Leu Asp Gln Leu His Gln Leu Arg Pro Ile
            500                 505                 510

Gly Thr Arg Gly Leu Lys Leu Pro Ser Leu Ser His Thr
        515                 520                 525
```

<210> 17
<211> 485
<212> PRT
<213> Homo sapiens

<400> 17
Met Ala Glu Leu Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro
1 5 10 15

Ala Pro Gly Pro Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro
20 25 30

Gly Leu His Ala Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln
35 40 45

Pro Asn Pro Leu Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly
50 55 60

Pro Asp Pro Leu Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro
65 70 75 80

Ser Ala Asn Ile Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser
85 90 95

Asp Leu Tyr Gly Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly
100 105 110

Pro Ser Gly Phe Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr
115 120 125

Gly Glu Ser Ala Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu
130 135 140

Ala Arg Tyr Val Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His
145 150 155 160

Val Ser Trp His Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His
165 170 175

Met Leu Leu Thr Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe
180 185 190

Gly Val Val Thr Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu
195 200 205

His Ser Ala Gln Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg

```
    210                 215                 220

Thr Val Pro Ile Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg
225                 230                 235                 240

Gly Glu Thr Ile Phe Glu Ile Asp Pro His Leu Gln Gln Glu Arg Val
                245                 250                 255

Asp Lys Gly Ile Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala
            260                 265                 270

Lys Cys Ala Trp Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp
        275                 280                 285

Ser Arg Ser Ile Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys
    290                 295                 300

Asp Thr Glu Gln Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn
305                 310                 315                 320

Ser Lys Pro Val Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu
                325                 330                 335

Pro His Asp Arg Ala Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser
            340                 345                 350

Ser Thr Ala Ile Ile Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu
        355                 360                 365

Ser Val His Leu Lys Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu
    370                 375                 380

Cys Leu Arg Gly Arg Leu Leu His Gly Arg His Phe Thr Tyr Lys Ser
385                 390                 395                 400

Ile Thr Gly Asp Met Ala Ile Thr Phe Val Ser Thr Gly Val Glu Gly
                405                 410                 415

Ala Phe Ala Thr Glu Glu His Pro Tyr Ala Ala His Gly Pro Trp Leu
            420                 425                 430

Gln Ile Leu Leu Thr Glu Glu Phe Val Glu Lys Met Leu Glu Asp Leu
        435                 440                 445

Glu Asp Leu Thr Ser Pro Glu Glu Phe Lys Leu Pro Lys Glu Tyr Ser
    450                 455                 460

Trp Pro Glu Lys Lys Leu Lys Val Ser Ile Leu Pro Asp Val Val Phe
```

```
465                 470                 475                 480

Asp Ser Pro Leu His
                485
```

<210> 18
<211> 389
<212> PRT
<213> Homo sapiens

<400> 18

```
Met Ala Glu Leu Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro
  1               5                  10                  15

Ala Pro Gly Pro Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro
             20                  25                  30

Gly Leu His Ala Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln
         35                  40                  45

Pro Asn Pro Leu Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly
     50                  55                  60

Pro Asp Pro Leu Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro
 65                  70                  75                  80

Ser Ala Asn Ile Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser
                 85                  90                  95

Asp Leu Tyr Gly Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly
            100                 105                 110

Pro Ser Gly Phe Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr
        115                 120                 125

Gly Glu Ser Ala Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu
    130                 135                 140

Ala Arg Tyr Val Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His
145                 150                 155                 160

Val Ser Trp His Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His
                165                 170                 175

Met Leu Leu Thr Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe
            180                 185                 190
```

```
Gly Val Val Thr Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu
        195                 200                 205

His Ser Ala Gln Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg
    210                 215                 220

Thr Val Pro Ile Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg
225                 230                 235                 240

Gly Glu Thr Ile Phe Glu Ile Asp Pro His Leu Gln Gln Glu Arg Val
                245                 250                 255

Asp Lys Gly Ile Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala
            260                 265                 270

Lys Cys Ala Trp Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp
        275                 280                 285

Ser Arg Ser Ile Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys
    290                 295                 300

Asp Thr Glu Gln Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn
305                 310                 315                 320

Ser Lys Pro Val Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu
                325                 330                 335

Pro His Asp Arg Ala Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser
            340                 345                 350

Ser Thr Ala Ile Ile Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu
        355                 360                 365

Ser Val His Leu Lys Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu
    370                 375                 380

Cys Leu Arg Phe Cys
385
```

<210> 19
<211> 360
<212> PRT
<213> Homo sapiens

<400> 19

```
Met Ala Glu Leu Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro
  1               5                  10                  15
```

```
Ala Pro Gly Pro Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro
            20                  25                  30

Gly Leu His Ala Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln
        35                  40                  45

Pro Asn Pro Leu Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly
    50                  55                  60

Pro Asp Pro Leu Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro
65                  70                  75                  80

Ser Ala Asn Ile Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser
                85                  90                  95

Asp Leu Tyr Gly Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly
            100                 105                 110

Pro Ser Gly Phe Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr
        115                 120                 125

Gly Glu Ser Ala Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu
    130                 135                 140

Ala Arg Tyr Val Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His
145                 150                 155                 160

Val Ser Trp His Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His
                165                 170                 175

Met Leu Leu Thr Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe
            180                 185                 190

Gly Val Val Thr Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu
        195                 200                 205

His Ser Ala Gln Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg
    210                 215                 220

Thr Val Pro Ile Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg
225                 230                 235                 240

Gly Glu Thr Ile Phe Glu Ile Asp Pro His Leu Gln Gln Glu Arg Val
                245                 250                 255

Asp Lys Gly Ile Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala
            260                 265                 270
```

```
Lys Cys Ala Trp Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp
        275                 280                 285

Ser Arg Ser Ile Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys
    290                 295                 300

Asp Thr Glu Gln Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn
305                 310                 315                 320

Ser Lys Pro Val Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu
                325                 330                 335

Pro His Asp Arg Ala Pro Asp Glu Val Thr Ala Ile Val Asn Thr Ser
            340                 345                 350

Thr Ser Glu Phe Leu Ala Gly Ala
        355                 360
```

<210> 20
<211> 497
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:myc-tagged SUFU-XL

<400> 20

```
Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met Ala Glu Leu
  1               5                  10                  15

Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro Ala Pro Gly Pro
             20                  25                  30

Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro Gly Leu His Ala
         35                  40                  45

Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln Pro Asn Pro Leu
     50                  55                  60

Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly Pro Asp Pro Leu
 65                  70                  75                  80

Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro Ser Ala Asn Ile
                 85                  90                  95
```

```
Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser Asp Leu Tyr Gly
            100                 105                 110

Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly Pro Ser Gly Phe
        115                 120                 125

Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr Gly Glu Ser Ala
    130                 135                 140

Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu Ala Arg Tyr Val
145                 150                 155                 160

Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His Val Ser Trp His
                165                 170                 175

Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His Met Leu Leu Thr
            180                 185                 190

Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe Gly Val Val Thr
        195                 200                 205

Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu His Ser Ala Gln
    210                 215                 220

Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg Thr Val Pro Ile
225                 230                 235                 240

Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg Gly Glu Thr Ile
                245                 250                 255

Phe Glu Ile Asp Pro His Leu Gln Gln Glu Arg Val Asp Lys Gly Ile
            260                 265                 270

Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala Lys Cys Ala Trp
        275                 280                 285

Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp Ser Arg Ser Ile
    290                 295                 300

Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys Asp Thr Glu Gln
305                 310                 315                 320

Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn Ser Lys Pro Val
                325                 330                 335

Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu Pro His Asp Arg
            340                 345                 350
```

```
Ala Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser Ser Thr Ala Ile
        355                 360                 365

Ile Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu Ser Val His Leu
    370                 375                 380

Lys Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu Cys Leu Arg Gly
385                 390                 395                 400

Arg Leu Leu His Gly Arg His Phe Thr Tyr Lys Ser Ile Thr Gly Asp
                405                 410                 415

Met Ala Ile Thr Phe Val Ser Thr Gly Val Glu Gly Ala Phe Ala Thr
            420                 425                 430

Glu Glu His Pro Tyr Ala Ala His Gly Pro Trp Leu Gln Ile Leu Leu
        435                 440                 445

Thr Glu Glu Phe Val Glu Lys Met Leu Glu Asp Leu Glu Asp Leu Thr
    450                 455                 460

Ser Pro Glu Glu Phe Lys Leu Pro Lys Glu Tyr Ser Trp Pro Glu Lys
465                 470                 475                 480

Lys Leu Lys Val Ser Ile Leu Pro Asp Val Val Phe Asp Ser Pro Leu
                485                 490                 495

His
```

<210> 21
<211> 401
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:myc-tagged
sufu-LK

<400> 21

```
Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met Ala Glu Leu
  1               5                  10                  15

Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro Ala Pro Gly Pro
            20                  25                  30
```

```
Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro Gly Leu His Ala
        35                  40                  45

Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln Pro Asn Pro Leu
    50                  55                  60

Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly Pro Asp Pro Leu
65                  70                  75                  80

Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro Ser Ala Asn Ile
                85                  90                  95

Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser Asp Leu Tyr Gly
            100                 105                 110

Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly Pro Ser Gly Phe
        115                 120                 125

Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr Gly Glu Ser Ala
    130                 135                 140

Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu Ala Arg Tyr Val
145                 150                 155                 160

Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His Val Ser Trp His
                165                 170                 175

Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His Met Leu Leu Thr
            180                 185                 190

Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe Gly Val Val Thr
        195                 200                 205

Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu His Ser Ala Gln
    210                 215                 220

Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg Thr Val Pro Ile
225                 230                 235                 240

Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg Gly Glu Thr Ile
                245                 250                 255

Phe Glu Ile Asp Pro His Leu Gln Gln Glu Arg Val Asp Lys Gly Ile
            260                 265                 270

Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala Lys Cys Ala Trp
        275                 280                 285
```

```
Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp Ser Arg Ser Ile
    290                 295                 300

Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys Asp Thr Glu Gln
305                 310                 315                 320

Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn Ser Lys Pro Val
                325                 330                 335

Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu Pro His Asp Arg
            340                 345                 350

Ala Pro Ser Arg Lys Asp Ser Leu Glu Ser Asp Ser Ser Thr Ala Ile
        355                 360                 365

Ile Pro His Glu Leu Ile Arg Thr Arg Gln Leu Glu Ser Val His Leu
    370                 375                 380

Lys Phe Asn Gln Glu Ser Gly Ala Leu Ile Pro Leu Cys Leu Arg Phe
385                 390                 395                 400

Cys
```

<210> 22
<211> 372
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:myc-tagged
SUFU-TT

<400> 22

```
Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Met Ala Glu Leu
  1               5                  10                  15

Arg Pro Ser Gly Ala Pro Gly Pro Thr Ala Pro Pro Ala Pro Gly Pro
            20                  25                  30

Thr Ala Pro Pro Ala Phe Ala Ser Leu Phe Pro Pro Gly Leu His Ala
        35                  40                  45

Ile Tyr Gly Glu Cys Arg Arg Leu Tyr Pro Asp Gln Pro Asn Pro Leu
    50                  55                  60

Gln Val Thr Ala Ile Val Lys Tyr Trp Leu Gly Gly Pro Asp Pro Leu
```

65 70 75 80

Asp Tyr Val Ser Met Tyr Arg Asn Val Gly Ser Pro Ser Ala Asn Ile
85 90 95

Pro Glu His Trp His Tyr Ile Ser Phe Gly Leu Ser Asp Leu Tyr Gly
100 105 110

Asp Asn Arg Val His Glu Phe Thr Gly Thr Asp Gly Pro Ser Gly Phe
115 120 125

Gly Phe Glu Leu Thr Phe Arg Leu Lys Arg Glu Thr Gly Glu Ser Ala
130 135 140

Pro Pro Thr Trp Pro Ala Glu Leu Met Gln Gly Leu Ala Arg Tyr Val
145 150 155 160

Phe Gln Ser Glu Asn Thr Phe Cys Ser Gly Asp His Val Ser Trp His
165 170 175

Ser Pro Leu Asp Asn Ser Glu Ser Arg Ile Gln His Met Leu Leu Thr
180 185 190

Glu Asp Pro Gln Met Gln Pro Val Gln Thr Pro Phe Gly Val Val Thr
195 200 205

Phe Leu Gln Ile Val Gly Val Cys Thr Glu Glu Leu His Ser Ala Gln
210 215 220

Gln Trp Asn Gly Gln Gly Ile Leu Glu Leu Leu Arg Thr Val Pro Ile
225 230 235 240

Ala Gly Gly Pro Trp Leu Ile Thr Asp Met Arg Arg Gly Glu Thr Ile
245 250 255

Phe Glu Ile Asp Pro His Leu Gln Gln Glu Arg Val Asp Lys Gly Ile
260 265 270

Glu Thr Asp Gly Ser Asn Leu Ser Gly Val Ser Ala Lys Cys Ala Trp
275 280 285

Asp Asp Leu Ser Arg Pro Pro Glu Asp Asp Glu Asp Ser Arg Ser Ile
290 295 300

Cys Ile Gly Thr Gln Pro Arg Arg Leu Ser Gly Lys Asp Thr Glu Gln
305 310 315 320

Ile Arg Glu Thr Leu Arg Arg Gly Leu Glu Ile Asn Ser Lys Pro Val

325 330 335

Leu Pro Pro Ile Asn Pro Gln Arg Gln Asn Gly Leu Pro His Asp Arg
340 345 350

Ala Pro Asp Glu Val Thr Ala Ile Val Asn Thr Ser Thr Ser Glu Phe
355 360 365

Leu Ala Gly Ala
370

<210> 23
<211> 1455
<212> DNA
<213> Homo sapiens

<400> 23
atggcggagc tgcggcctag cggcgccccc ggccccaccg cgcccccggc ccctggcccg 60
actgcccccc cggctttcgc ttcgctcttt cccccgggac tgcacgccat ctacggagag 120
tgccgccgcc tttaccctga ccagccgaac ccgctccagg ttaccgctat cgtcaagtac 180
tggttgggtg gcccagaccc cttggactat gttagcatgt acaggaatgt ggggagccct 240
tctgctaaca tccccgagca ctggcactac atcagcttcg gcctgagtga tctctatggt 300
gacaacagag tccatgagtt tacaggaaca gatggaccta gtggttttgg ctttgagttg 360
acctttcgtc tgaagagaga aactggggag tctgccccac caacatggcc cgcagagtta 420
atgcagggct tggcacgata cgtgttccag tcagagaaca ccttctgcag tggggaccat 480
gtgtcctggc acagcccttt ggataacagt gagtcaagaa ttcagcacat gctgctgaca 540
gaggacccac agatgcagcc cgtgcagaca ccctttgggg tagttacctt cctccagatc 600
gttggtgtct gcactgaaga gctacactca gcccagcagt ggaacgggca gggcatcctg 660
gagctgctgc ggacagtgcc tattgctggc ggcccctggc tgataactga catgcggagg 720
ggagagacca tatttgagat cgatccacac ctgcaagaga gagttgacaa aggcatcgag 780
acagatggct ccaacctgag tggtgtcagt gccaagtgtg cctgggatga cctgagccgg 840
ccccccgagg atgacgagga cagccggagc atctgcatcg gcacacagcc ccggcgactc 900
tctggcaaag acacagagca gatccgggag accctgagga gaggactcga gatcaacagc 960
aaacctgtcc ttccaccaat caaccctcag cggcagaatg gtctccccca cgaccgggcc 1020
ccgagccgca aagacagcct ggaaagtgac agctccacgg ccatcattcc ccatgagctg 1080
attcgcacgc ggcagcttga gagcgtacat ctgaaattca accaggagtc cggagccctc 1140
attcctctct gcctaagggg caggctcctg catggacggc actttacata taaaagtatc 1200
acaggtgaca tggccatcac gtttgtctcc acgggagtgg aaggcgcctt tgccactgag 1260
gagcatcctt acgcggctca tggaccctgg ttacaaattc tgttgaccga agagtttgta 1320
gagaaaatgt tggaggattt agaagatttg acttctccag aggaattcaa acttcccaaa 1380
gagtacagct ggcctgaaaa gaagctgaag gtctccatcc tgcctgacgt ggtgttcgac 1440
agtccgctac actag 1455

<210> 24
<211> 1458
<212> DNA
<213> Homo sapiens

<400> 24
atggcggagc tgcggcctag cggcgccccc ggccccaccg cgcccccggc ccctggcccg 60
actgcccccc cggctttcgc ttcgctcttt cccccgggac tgcacgccat ctacggagag 120
tgccgccgcc tttaccctga ccagccgaac ccgctccagg ttaccgctat cgtcaagtac 180
tggttgggtg gcccagaccc cttggactat gttagcatgt acaggaatgt ggggagccct 240
tctgctaaca tccccgagca ctggcactac atcagcttcg gcctgagtga tctctatggt 300
gacaacagag tccatgagtt tacaggaaca gatggaccta gtggttttgg ctttgagttg 360
acctttcgtc tgaagagaga aactggggag tctgccccac caacatggcc cgcagagtta 420
atgcagggct tggcacgata cgtgttccag tcagagaaca ccttctgcag tggggaccat 480
gtgtcctggc acagcccttt ggataacagt gagtcaagaa ttcagcacat gctgctgaca 540
gaggacccac agatgcagcc cgtgcagaca ccctttgggg tagttacctt cctccagatc 600
gttggtgtct gcactgaaga gctacactca gcccagcagt ggaacgggca gggcatcctg 660
gagctgctgc ggacagtgcc tattgctggc ggcccctggc tgataactga catgcggagg 720
ggagagacca tatttgagat cgatccacac ctgcaacagg agagagttga caaaggcatc 780
gagacagatg gctccaacct gagtggtgtc agtgccaagt gtgcctggga tgacctgagc 840
cggcccccgg aggatgacga ggacagccgg agcatctgca tcggcacaca gccccggcga 900
ctctctggca aagacacaga gcagatccgg gagaccctga ggagaggact cgagatcaac 960
agcaaacctg tccttccacc aatcaaccct cagcggcaga atggtctccc ccacgaccgg 1020
gccccgagcc gcaaagacag cctggaaagt gacagctcca cggccatcat tccccatgag 1080
ctgattcgca cgcggcagct tgagagcgta catctgaaat tcaaccagga gtccggagcc 1140
ctcattcctc tctgcctaag gggcaggctc ctgcatggac ggcactttac atataaaagt 1200
atcacaggtg acatggccat cacgtttgtc tccacgggag tggaaggcgc ctttgccact 1260
gaggagcatc cttacgcggc tcatggaccc tggttacaaa ttctgttgac cgaagagttt 1320
gtagagaaaa tgttggagga tttagaagat ttgacttctc cagaggaatt caaacttccc 1380
aaagagtaca gctggcctga aaagaagctg aaggtctcca tcctgcctga cgtggtgttc 1440
gacagtccgc tacactag 1458

<210> 25
<211> 1319
<212> DNA
<213> Homo sapiens

<400> 25
atggcggagc tgcggcctag cggcgccccc ggccccaccg cgcccccggc ccctggcccg 60
actgcccccc cggctttcgc ttcgctcttt cccccgggac tgcacgccat ctacggagag 120
tgccgccgcc tttaccctga ccagccgaac ccgctccagg ttaccgctat cgtcaagtac 180
tggttgggtg gcccagaccc cttggactat gttagcatgt acaggaatgt ggggagccct 240
tctgctaaca tccccgagca ctggcactac atcagcttcg gcctgagtga tctctatggt 300
gacaacagag tccatgagtt tacaggaaca gatggaccta gtggttttgg ctttgagttg 360
acctttcgtc tgaagagaga aactggggag tctgccccac caacatggcc cgcagagtta 420
atgcagggct tggcacgata cgtgttccag tcagagaaca ccttctgcag tggggaccat 480
gtgtcctggc acagcccttt ggataacagt gagtcaagaa ttcagcacat gctgctgaca 540
gaggacccac agatgcagcc cgtgcagaca ccctttgggg tagttacctt cctccagatc 600
gttggtgtct gcactgaaga gctacactca gcccagcagt ggaacgggca gggcatcctg 660
gagctgctgc ggacagtgcc tattgctggc ggcccctggc tgataactga catgcggagg 720
ggagagacca tatttgagat cgatccacac ctgcaacagg agagagttga caaaggcatc 780
gagacagatg gctccaacct gagtggtgtc agtgccaagt gtgcctggga tgacctgagc 840

cggccccccg aggatgacga ggacagccgg agcatctgca tcggcacaca gccccggcga 900
ctctctggca aagacacaga gcagatccgg gagaccctga ggagaggact cgagatcaac 960
agcaaacctg tccttccacc aatcaaccct cagcggcaga atggtctccc ccacgaccgg 1020
gccccgagcc gcaaagacag cctggaaagt gacagctcca cggccatcat tccccatgag 1080
ctgattcgca cgcggcagct tgagagcgta catctgaaat tcaaccagga gtccggagcc 1140
ctcattcctc tctgcctaag attctgttga ccgaagagtt tgtagagaaa atgttggagg 1200
atttagaaga tttgacttct ccagaggaat tcaaacttcc caaagagtac agctggcctg 1260
aaaagaagct gaaggtctcc atcctgcctg acgtggtgtt cgacagtccg ctacactag 1319

<210> 26
<211> 1657
<212> DNA
<213> Homo sapiens

<400> 26
atggcggagc tgcggcctag cggcgccccc ggccccaccg cgcccccggc ccctggcccg 60
actgcccccc cggctttcgc ttcgctcttt cccccgggac tgcacgccat ctacggagag 120
tgccgccgcc tttaccctga ccagccgaac ccgctccagg ttaccgctat cgtcaagtac 180
tggttgggtg gcccagaccc cttggactat gttagcatgt acaggaatgt ggggagccct 240
tctgctaaca tccccgagca ctggcactac atcagcttcg gcctgagtga tctctatggt 300
gacaacagag tccatgagtt tacaggaaca gatggaccta gtggttttgg ctttgagttg 360
acctttcgtc tgaagagaga aactggggag tctgccccac caacatggcc cgcagagtta 420
atgcagggct tggcacgata cgtgttccag tcagagaaca ccttctgcag tggggaccat 480
gtgtcctggc acagcccttt ggataacagt gagtcaagaa ttcagcacat gctgctgaca 540
gaggacccac agatgcagcc cgtgcagaca ccctttgggg tagttacctt cctccagatc 600
gttggtgtct gcactgaaga gctacactca gcccagcagt ggaacgggca gggcatcctg 660
gagctgctgc ggacagtgcc tattgctggc ggcccctggc tgataactga catgcggagg 720
ggagagacca tatttgagat cgatccacac ctgcaacagg agagagttga caaaggcatc 780
gagacagatg gctccaacct gagtggtgtc agtgccaagt gtgcctggga tgacctgagc 840
cggccccccg aggatgacga ggacagccgg agcatctgca tcggcacaca gccccggcga 900
ctctctggca aagacacaga gcagatccgg gagaccctga ggagaggact cgagatcaac 960
agcaaacctg tccttccacc aatcaaccct cagcggcaga atggtctccc ccacgaccgg 1020
gccccagatg aagtcacggc cattgttaac acctcaactt cagaattcct tgctggtgct 1080
tgaattctca gtccaaacgc tgcatccccc tccctcgagt tgcacaatca ggggtatttt 1140
tatcgcccat ccttgatcct cggagtcacc caacacttca gcagtgagtc acttctccag 1200
ggaagggaaa aggaaggctc aaaggagccg caaagacagc ctggaaagtg acagctccac 1260
ggccatcatt ccccatgagc tgattcgcac gcggcagctt gagagcgtac atctgaaatt 1320
caaccaggag tccggagccc tcattcctct ctgcctaagg ggcaggctcc tgcatggacg 1380
gcactttaca tataaaagta tcacaggtga catggccatc acgtttgtct ccacgggagt 1440
ggaaggcgcc tttgccactg aggagcatcc ttacgcggct catggaccct ggttacaaat 1500
tctgttgacc gaagagtttg tagagaaaat gttggaggat ttagaagatt tgacttctcc 1560
agaggaattc aaacttccca aagagtacag ctggcctgaa aagaagctga aggtctccat 1620
cctgcctgac gtggtgttcg acagtccgct acactag 1657

<210> 27
<211> 3326
<212> DNA
<213> Homo sapiens

<400> 27
atgttcaact cgatgacccc accaccaatc agtagctatg gcgagccctg ctgtctccgg 60
cccctcccca gtcagggggc ccccagtgtg gggacagaag gactgtctgg cccgcccttc 120
tgccaccaag ctaacctcat gtccggcccc cacagttatg ggccagccag agagaccaac 180
agctgcaccg agggcccact cttttcttct ccccggagtg cagtcaagtt gaccaagaag 240
cgggcactgt ccatctcacc tctgtcggat gccagcctgg acctgcagac ggttatccgc 300
acctcaccca gctccctcgt agctttcatc aactcgcgat gcacatctcc aggaggctcc 360
tacggtcatc tctccattgg caccatgagc ccatctctgg gattcccagc ccagatgaat 420
caccaaaaag ggccctcgcc ttcctttggg gtccagcctt gtggtcccca tgactctgcc 480
cggggtggga tgatcccaca tcctcagtcc cggggaccct tcccaacttg ccagctgaag 540
tctgagctgg acatgctggt tggcaagtgc cgggaggaac ccttggaagg tgatatgtcc 600
agccccaact ccacaggcat acaggatccc ctgttgggga tgctggatgg gcgggaggac 660
ctcgagagag aggagaagcg tgagcctgaa tctgtgtatg aaactgactg ccgttgggat 720
ggctgcagcc aggaatttga ctcccaagag cagctggtgc accacatcaa cagcgagcac 780
atccacgggg agcggaagga gttcgtgtgc cactgggggg gctgctccag ggagctgagg 840
cccttcaaag cccagtacat gctggtggtt cacatgcgca gacacactgg cgagaagcca 900
cacaagtgca cgtttgaagg gtgccggaag tcatactcac gcctcgaaaa cctgaagacg 960
cacctgcggt cacacacggg tgagaagcca tacatgtgtg agcacgaggg ctgcagtaaa 1020
gccttcagca atgccagtga ccgagccaag caccagaatc ggacccattc caatgagaag 1080
ccgtatgtat gtaagctccc tggctgcacc aaacgctata cagatcctag ctcgctgcga 1140
aaacatgtca agacagtgca tggtcctgac gcccatgtga ccaaacggca ccgtggggat 1200
ggcccccctgc ctcgggcacc atccatttct acagtggagc ccaagaggga gcgggaagga 1260
ggtcccatca gggaggaaag cagactgact gtgccagagg gtgccatgaa gccacagcca 1320
agccctgggg cccagtcatc ctgcagcagt gaccactccc cggcagggag tgcagccaat 1380
acagacagtg gtgtggaaat gactggcaat gcaggggcca gcactgaaga cctctccagc 1440
ttggacgagg gaccttgcat tgctggcact ggtctgtcca ctcttcgccg ccttgagaac 1500
ctcaggctgg accagctaca tcaactccgg ccaataggga cccggggtct caaactgccc 1560
agcttgtccc acaccggtac cactgtgtcc cgccgcgtgg gccccccagt ctctcttgaa 1620
cgccgcagca gcagctccag cagcatcagc tctgcctata ctgtcagccg ccgctcctcc 1680
ctggcctctc ctttcccccc tggctcccca ccagagaatg gagcatcctc cctgcctggc 1740
cttatgcctg cccagcacta cctgcttcgg gcaagatatg cttcagccag agggggtggt 1800
acttcgccca ctgcagcatc cagcctggat cggataggtg gtcttcccat gcctccttgg 1860
agaagccgag ccgagtatcc aggatacaac cccaatgcag gggtcacccg gagggccagt 1920
gacccagccc aggctgctga ccgtcctgct ccagctagag tccagaggtt caagagcctg 1980
ggctgtgtcc ataccccacc cactgtggca gggggaggac agaactttga tccttacctc 2040
ccaacctctg tctactcacc acagcccccc agcatcactg agaatgctgc catggatgct 2100
agagggctac aggaagagcc agaagttggg acctccatgg tgggcagtgg tctgaacccc 2160
tatatggact tcccacctac tgatactctg ggatatgggg gacctgaagg ggcagcagct 2220
gagccttatg gagcgagggg tccaggctct ctgcctcttg ggcctggtcc acccaccaac 2280
tatggcccca acccctgtcc ccagcaggcc tcatatcctg accccaccca agaaacatgg 2340
ggtgagttcc cttcccactc tgggctgtac ccaggcccca aggctctagg tggaacctac 2400
agccagtgtc ctcgacttga acattatgga caagtgcaag tcaagccaga acaggggtgc 2460
ccagtggggt ctgactccac aggactggca ccctgcctca atgcccaccc cagtgagggg 2520
cccccacatc cacagcctct cttttcccat tacccccagc cctctcctcc ccaatatctc 2580
cagtcaggcc cctataccca gccacccccct gattatcttc cttcagaacc caggccttgc 2640
ctggactttg attcccccac ccattccaca gggcagctca aggctcagct tgtgtgtaat 2700
tatgttcaat ctcaacagga gctactgtgg gagggtgggg gcagggaaga tgcccccgcc 2760

caggaacctt cctaccagag tcccaagttt ctgggggttt cccaggttag cccaagccgt 2820
gctaaagctc cagtgaacac atatggacct ggctttggac ccaacttgcc caatcacaag 2880
tcaggttcct atcccacccc ttcaccatgc catgaaaatt ttgtagtggg ggcaaatagg 2940
gcttcacata gggcagcagc accacctcga cttctgcccc cattgcccac ttgctatggg 3000
cctctcaaag tgggaggcac aaaccccagc tgtggtcatc ctgaggtggg caggctagga 3060
gggggtcctg ccttgtaccc tcctcccgaa ggacaggtat gtaacccccc ggactctctt 3120
gatcttgaca acactcagct ggactttgtg gctattctgg atgagcccca ggggctgagt 3180
cctcctcctt cccatgatca gcggggcagc tctggacata ccccacctcc ctctgggccc 3240
cccaacatgg ctgtgggcaa catgagtgtc ttactgagat ccctacctgg ggaaacagaa 3300
ttcctcaact ctagtgccta gtcgac 3326

<210> 28
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide hybridising with the SUFUH gene

<400> 28
cagttgtgtc aacgagatct cc 22

<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide hybridising with the human SUFUH gene

<400> 29
ctgtggcctg tgcatggcac 20

<210> 30
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide hybridising with the human SUFUH gene

<400> 30
ccgctggcta agccttgtga c 21

<210> 31

<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Oligonucleotide hybridising with the human SUFUH gene

<400> 31
cagggtccag agtcaaacct ca 22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 00/01576

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C07K 7/04, A61K 38/10, A61P 35/00
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C07K, A61K, A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. Mol Med, Volume 77, 1999, Heidi Hahn et al, "The patched signaling pathway in tumorigenesis and development: lessons from animal models" page 459 - page 468 --	1-13
A	WO 9932517 A1 (KAROLINSKA INNOVATIONS AB), 1 July 1999 (01.07.99) -- --------	1-13

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
1 December 2000	11 -12- 2000
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Carolina Gómez Lagerlöf/EÖ Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Information on patent family members

02/11/00

International application No.
PCT/SE 00/01576

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9932517 A1	01/07/99	AU 1991799 A	12/07/99
		AU 1991899 A	12/07/99
		EP 1037920 A	27/09/00
		SE 9704788 D	00/00/00
		WO 9932609 A	01/07/99
		AU 1795699 A	05/07/99
		EP 1042762 A	11/10/00
		SE 9802293 D	00/00/00